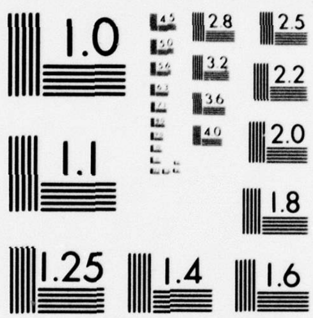
AD-A060 828 WEAPONS RESEARCH ESTABLISHMENT SALISBURY (AUSTRALIA) F/G 9/4
ADAPTIVE BINARY CODING SCHEMES FOR TRADING DATA RATE FOR REDUCE--ETC(U)
FEB 78 J HAYWARD
UNCLASSIFIED WRE-TR-1910(W) NL

1 OF 1
AD
A060828

END
DATE
FILMED
01-79
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

14 WRE-TR-1910(W)

12

AR-001-153

AD A060828

DDC FILE COPY

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

WEAPONS RESEARCH ESTABLISHMENT

SALISBURY, SOUTH AUSTRALIA

9 TECHNICAL REPORT 1910 (W)

6

ADAPTIVE BINARY CODING SCHEMES FOR TRADING DATA RATE FOR REDUCED SIGNAL TO NOISE RATIO REQUIREMENT

10 J. HAYWARD

Approved for Public Release

12 93p.

COPY No. 21

11 FEBR[redacted]1978

78 11 02 044

371 700

JOB

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

UNCLASSIFIED

LEVEL II

12

AR-001-153

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

WEAPONS RESEARCH ESTABLISHMENT

TECHNICAL REPORT 1910 (W)

ADAPTIVE BINARY CODING SCHEMES FOR TRADING DATA RATE FOR REDUCED SIGNAL TO NOISE RATIO REQUIREMENT

J. Hayward

S U M M A R Y

A detailed analysis is made of methods of increasing the redundancy of binary level data words so as to permit transmission at a reduced signal to noise ratio without an increase in error probability. Repeated transmission of the same data word is shown to be rather inefficient, especially if the integration is carried out after stages of non-linear detection and/or decoding. The need for care in ensuring the statistical independence between replicates, and between the individual bits of longer codewords, is shown. A cascade of a Golay cyclical error correcting coding process, followed by an orthogonal sparse coding, using maximum amplitude detection instead of threshold detection, is found to provide an efficient system where signal to noise ratio is restricted and energy consumption at the transmitter must be minimised. Replication can be used to provide an additional trade off during periods of abnormally low signal to noise ratio. An analysis is also given to enable the calculation of losses arising from commonly used quasi-matched filters for the extraction of signal elements from Gaussian noise.

Approved for Public Release

POSTAL ADDRESS: The Director, Weapons Research Establishment, Box 2151, G.P.O., Adelaide, South Australia, 5001.

UNCLASSIFIED

DOCUMENT CONTROL DATA SHEET

Security classification of this page: UNCLASSIFIED

1 DOCUMENT NUMBERS

AR Number:	AR-001-153
Report Number:	WRE-TR-1910(W)
Other Numbers:	

2 SECURITY CLASSIFICATION

a. Complete Document:	Unclassified
b. Title in Isolation:	Unclassified
c. Summary in Isolation:	Unclassified

3 TITLE

ADAPTIVE BINARY CODING SCHEMES FOR TRADING DATA RATE FOR REDUCED SIGNAL TO NOISE RATIO REQUIREMENTS

4 PERSONAL AUTHOR(S):

J. Hayward

5 DOCUMENT DATE:

February 1978

6

6.1 TOTAL NUMBER OF PAGES	92
6.2 NUMBER OF REFERENCES:	11

7 7.1 CORPORATE AUTHOR(S):

Weapons Research Establishment

7.2 DOCUMENT (WING) SERIES AND NUMBER

Weapons Research & Development Wing
TR-1910

8 REFERENCE NUMBERS

a. Task:	74/03
b. Sponsoring Agency:	RD73

9 COST CODE:

323/346762

10 IMPRINT (Publishing establishment):

Weapons Research Establishment

11 COMPUTER PROGRAM(S)
(Title(s) and language(s))

12 RELEASE LIMITATIONS (of the document):

Approved for Public Release

12.0	OVERSEAS	NO		P.R.	1	A		B		C		D		E	

Security classification of this page: UNCLASSIFIED

Security classification of this page: UNCLASSIFIED

13 ANNOUNCEMENT LIMITATIONS (of the information on these pages):

No limitation.

14 DESCRIPTORS:			15 COSATI CODES:
a. EJC Thesaurus Terms	Coding Decoding Error correction codes Coding theory Binary digits Signal to noise ratio	Filters Error analysis	0904 1201
b. Non-Thesaurus Terms	Golay error correction		

16 LIBRARY LOCATION CODES (for libraries listed in the distribution):

SW SR SD AACA

17 SUMMARY OR ABSTRACT:
(if this is security classified, the announcement of this report will be similarly classified)

A detailed analysis is made of methods of increasing the redundancy of binary level data words so as to permit transmission at a reduced signal to noise ratio without an increase in error probability. Repeated transmission of the same data word is shown to be rather inefficient, especially if the integration is carried out after stages of non-linear detection and/or decoding. The need for care in ensuring the statistical independence between replicates, and between the individual bits of longer codewords, is shown. A cascade of a Golay cyclical error correcting coding process, followed by an orthogonal sparse coding, using maximum amplitude detection instead of threshold detection, is found to provide an efficient system where signal to noise ratio is restricted and energy consumption at the transmitter must be minimised. Replication can be used to provide an additional trade off during periods of abnormally low signal to noise ratio. An analysis is also given to enable the calculation of losses arising from commonly used quasi-matched filters for the extraction of signal elements from Gaussian noise.

Security classification of this page: UNCLASSIFIED

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

12. Word error probability. One out of (K + 1) code. Element Threshold ($P_A = P_B$)

13. Integration loss. Maximum amplitude decoding

14. Post detector integration performance (n replicates). One of (K + 1) code. Maximum amplitude detector

15. Post detector integration performance (n replicates). (K + 1) code. Summation of element threshold crossings

16. Golay code error probability with post decoder integration.

17. Cascaded (K + 1) and Golay performance. No interleaving of bits

18. Cascaded (K + 1) and Golay performance. Full interleaving

19. Cascaded codes. Within group replication. Random bit allocation between Golay words

20. Replication of 12-bit data word

21. Unreplicated 12-bit word via (K + 1) code

22. Candidate coding schemes

23. Cascaded Golay and (K + 1) code. (K + 1) = 4. Scheme 6

24. Quasi-matched filter schemes for individual elements

1. INTRODUCTION

The signalling of data between two stations, separated in space, involves the manipulation and observation of some physical parameter, available at both sites (e.g. the voltage in a telephone cable, or the field strength of an electromagnetic wave in space). Providing quantum effects are negligible, most physical parameters can be varied continuously in both magnitude and time. They are also subject to naturally occurring variations (noise), due to thermal excitation of the constituent molecules of the medium, as well as to bursts of larger variations caused by other natural, or man-made processes. In some situations, variations may be man-made with the deliberate intention of interfering in the data transfer.

Most man-made data is discrete in both magnitude and time; being a sequence of time samples of some conceptual variable (which may or may not be related to a continuous physical variable) with a finite number of possible levels (determined in the absence of other indications, by the number of significant figures in its representation). During the transfer of data, it is usual for the manipulated variable to be held fixed at some discrete level, for some short element of time; with more or less rapid change to another level at the beginning of the next element. The observer's task is therefore to average the magnitude of the received variable over the duration of each time element, and to choose between the predetermined possible levels, at the end of each element. If the levels are separated sufficiently, compared with the expected variations in the time average due to noise, the basic error probability will be low.

Design of a data link, therefore, involves the selection of a basic element length and set of levels of magnitude, together with an encoding plan, which provides a (usually one to one) relation between the data sequence and levels, and the element sequence of levels. The element length is necessarily a compromise between excessive transmission time, and excessive channel bandwidth. The number and separation of the basic levels is a compromise between the element error probability, and the peak power required to manipulate (modulate) the physical medium. The encoding method provides for a trade-off between data error probability, data transmission time, and the total energy required to transmit a segment of data.

Shannon, and others, have derived the well known optimum limit to the trade-off relationship between the data rate, bandwidth, signal to noise ratio, and error probability for a one-way channel. This requires a physical transmission parameter, unbounded in both directions, and a very complex coding and decoding process. Practical constraints usually result in a transmission scheme departing considerably from this optimum and binary level coding is commonly used. If the signal to noise ratio is also a variable (perhaps due to changes in distance between the two stations, linked by a radio path) the best encoding and transmission scheme is not at all obvious. If a reduced data rate is tolerable under poor conditions of signal to noise ratio, an adaptive encoding scheme, trading off data rate and signal to noise ratio, may be appropriate. One method involves multiple transmission (replication) of the same data block. Another, more efficient method, involves cascaded encodings, switched in or out as required, to vary the redundancy of the data. As the complexity of the encoding and decoding processes, and the efficiency of the trade-off, vary widely between different methods, selection of an appropriate scheme is not an easy matter. This is compounded by the difficulty in calculating the trade-off efficiency.

This report is concerned with a detailed investigation of a rather special case, previously examined briefly by Sarkies(ref.9), in the hope that it will illustrate the methods of attack required; that it will give a feel for the results achievable in this, and other similar schemes; and that it will provide a wealth of data for the design of an adaptive data link with practical applications.

The basic element carries only two levels, namely the presence or absence of a fixed frequency carrier wave, with constant amplitude. The phase of the carrier is unpredictable, so coherent detection is not possible. The competition is restricted to additive Gaussian noise. A block of data to be transmitted, is assumed to consist of 12 bits of information, an error in one or more of which renders the entire block useless. No consideration is given to the use of a return channel for error correction purposes. Encoding schemes considered are:

(a) Direct encoding where each of the 12 data bits is carried by one of twelve successive elements. Each element may contain a pulse of carrier.

(b) One out of (K + 1) encoding (denoted orthogonal encoding by Sarkies), where (K + 1) successive elements contain one, and only one, pulse of carrier. The $\log_2(K + 1)$ bits of data from several groups are combined to form the 12-bit data word. Values of (K + 1) considered, are 2, 4, 8, and 16.

(c) Golay encoding, where the 12 data bits are combined with 11 redundant bits, according to the cyclical Golay code format, to form a 23 bit word for transmission via schemes (a) or (b) above.

(d) Replication of any of the above schemes. The multiple (n) takes values in the range 1 to 10.

The replicates can be combined (integrated) before carrier detection, after carrier detection but before element (or (K + 1) element group) decision, after element decision but before Golay decoding, or after the production of the individual 12 bit data words. Figure 22 illustrates the possible integration points. Sarkies considered only the second case, which is fairly efficient, but also fairly difficult to carry out in practice. The other methods have been evaluated here, to permit trade-off of efficiency and simplicity. In general, they vary inversely. It appears that each time the data is passed through a non-linear process (of which each decoding stage, and most carrier detection methods is one) the integration efficiency is reduced. In any case, replication is the least efficient (but easiest to carry out) of all the methods of trading off data rate for signal to noise ratio. It is important to ensure that replicates are assembled in such a way that the noise samples are statistically independent.

Section 2 considers the averaging of the received signal over one time element. The performance of the ideally matched filter is calculated and the relationship to several quasi-matched filters is identified. Figure 24 summarises the possibilities. The performance of linear envelope detection is compared with coherent detection when used ahead of a threshold detector for element decision. This provides a consistent basis on which the main sections of the report are built. Section 3 compares bi-phase modulation with on-off keying and is not part of the main thrust of the report.

Section 4 considers the integration of replicated elements, with the integration occurring before, between and after the carrier and threshold detectors.

Section 5 is concerned with the performance of un-replicated one out of (K + 1) coding, using both threshold and maximum amplitude element decision methods. The various methods of replication integration are considered in Section 6.

Section 7 introduces the Golay code and determines its performance, with and without replication integration. The use of correction weight to effect a decision between two dissimilar replicates, is introduced. This provides the only possibility of obtaining an integration improvement from only two replicates, with any type of post decoding integration. All others require a chance selection, which gives no improvement.

Section 8 considers the cascading of (K + 1) and Golay codes, with and without replication integration. Careful attention is required to ensure that replicates are not statistically dependent.

By this stage, sufficient information is available to analyse the performance of all possible combinations of coding and integration schemes. For comparison purposes, and to create some order, Section 9 presents an investigation of all non-replicated methods of achieving a 12-bit data word, and Table 12 summarises the data rates, signal to noise ratios and relative energy per data word, required to give a word error probability of 10^{-4}. The cascaded (K + 1) and Golay codes, using maximum amplitude detection, and (K + 1) = 4, is found to be the best choice for applications where transmitted carrier power is severely limited, and energy drain is important. The data rate is only one quarter of the maximum possible (direct encoded scheme), but the signal to noise ratio required is reduced by a very significant, 6.7 dB. If the peak transmitter power is reduced by this amount, there is a consequent reduction in the energy drain of 3.7 dB. This scheme is chosen for an examination of the improvement to be obtained by the various possible forms of replication integration. Figure 23 illustrates the results for a small number of replicates. For 2 replicates, and 10^{-4} probability of error, we observe signal to noise improvements of 3.0, 2.25, 0.0 and 1.4 dB for precarrier detector, pre-maximum amplitude detector, post-maximum amplitude detector, and post-Golay decoding, location of the replication integration. For 4 replicates, the figures become 6.0, 4.4, 2.95 and 2.6 dB, respectively.

2. ERROR PROBABILITY WITH BINARY AMPLITUDE MODULATION

2.1 General

In this section we consider the transmission of data, using two level amplitude modulation (binary) of a radio frequency carrier, and the errors to be expected in decoding the received signal in competition with white noise.

We assume that a signal element consists of a time interval, of length, τ, during which a pulse of energy is either present or absent, depending on the data transmitted over that element. Suppose the power received when the pulse is present is W watts, and the white noise power density, referred to the same point, is N watts per Hz (only positive frequencies being considered). We assume also, that the time epoch, t, at the start of the element, is known at the receiver. We wish to process the signal received over the period, τ, and make a decision, with low probability of error, as to whether the signal was present, or not.

North(ref.1) and others, have shown that, under the conditions postulated above, a matched filter produces a derived signal output with the maximum possible signal to noise ratio. If this signal is compared with a suitable threshold, a binary decision can be made as to the probable data bit value. The threshold would normally be chosen to equalise the probability of both types of error, i.e. in deciding that a signal is present when it is not, and vice versa.

2.2 Integrating matched filter

A matched filter can take several forms. The simplest to analyse, consists of a multiplication process with a coherent reference frequency, followed by integration of the result over the interval, τ. The derived signal is the value at the end of the interval. The analysis is carried out in Appendix I, where it is shown that the derived signal has a Gaussian distribution with a signal to noise ratio, γ_0, of $\frac{2\ W.\tau}{N}$, that is, twice the ratio of signal energy within the element, to the noise power density. Because of the process linearity, the noise distribution is unchanged when the signal is absent. Figure 1 shows the distributions, with and without

signals, and the location of the optimum threshold, which has the value $\sqrt{W}/2$.

The probability of either form of error is given by:

$$P_{e1} = \int_{\sqrt{W}/2}^{\infty} \frac{dP_e(x)}{dx} \,.\, dx = \int_{\sqrt{W}/2}^{\infty} \frac{1}{\sigma \,.\, \sqrt{2\pi}} \,.\, e^{-\frac{x^2}{2\sigma^2}} \,.\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\sqrt{W}/2\sigma}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \,.\, d(\frac{x}{\sigma})$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\sqrt{\frac{W.\tau}{2N}}}^{\infty} e^{-\frac{t^2}{2}} \,.\, dt \qquad (1)$$

where t is a dummy variable.

This is plotted in figure 2, with logarithmic scales, and with $X = \frac{W.\tau}{N}$.

Figure 3 shows the magnitude of the normalised optimum threshold, $(\sqrt{\frac{W.\tau}{N}}/2 = \sqrt{X}/2)$ also plotted as a function of X.

2.3 Tapped delay line matched filter

A difficulty in implementing the previous scheme of matched filter, is the need to know, in advance and with fair accuracy, the epoch, t, to begin the integration. Furthermore there is no indication afterwards, of an error in t. This difficulty is largely overcome with another form of matched filter, in which the received signal is fed into a delay line of length, τ. The signals at n taps, suitably spaced, are combined by summation, and the result is multiplied by a coherent reference frequency. No integration is required. (The delay line carries out this function, and responds to signals at all epochs, t). The derived signal is the output of the multiplier, low-pass filtered to remove multiples of the carrier/reference frequency. This scheme is analysed in Appendix II, where it is shown that the performance is the same as for the previous scheme. The tapped delay line produces a sin x/x amplitude versus frequency response about the reference frequency, which appropriately shapes the noise, while maintaining a unity gain response to the signal. Since the overall response is matched, and the coherent multiplication is linear, the signal to in-phase noise ratio at the output of the sin x/x filter is the same as that at the derived signal (i.e. $\frac{2\ W.\tau}{N}$, from Appendix I). Since the quadrature noise power equals the in-phase noise power, the signal to total noise power ratio at the output of the sin x/x filter is $X_1 = \frac{W.\tau}{N}$, giving an equivalent noise bandwidth of $\frac{1}{\tau}$ for the sin x/x filter.

2.4 Envelope detection

Both of the previous schemes suffer from the need to correlate the phases of the received and reference waves. This problem is avoided, at the penalty of reduced performance at lower signal to noise ratios, if a linear envelope detector is substituted, for the coherent multiplication, at the output of the sin x/x filter. The envelope, at time $(t + \tau)$ is given by the relation:

$$v_e = \sqrt{(\sqrt{2w} + x_n)^2 + y_n^2} \qquad (2)$$

where x_n and y_n are independent samples from a Gaussian distribution of zero mean, and variance $\frac{N}{\tau}$, and represent in-phase and quadrature noise components. Normalising with respect to the total r.m.s. noise, $\sigma_n = \sqrt{\frac{N}{\tau}}$:

$$v_o = \frac{v_e}{\sigma_n} = \sqrt{(\sqrt{2x} + \frac{x_n}{\sigma_n})^2 + (\frac{y_n}{\sigma_n})^2} \qquad (3)$$

where X is the total signal to noise ratio at the detector (i.e. $\frac{W.\tau}{N}$). Rice has given the following relations for probability density, mean and variance of this distribution (which is non-Gaussian):

$$\frac{dP(v_o)}{dv_o} = v_o \cdot e^{-\frac{v_o^2 + 2x}{2}} \cdot I_o(\sqrt{2x}.v_o) \qquad (4)$$

$$\bar{v}_o = \sqrt{\frac{\pi}{2}} \cdot e^{-\frac{X}{2}} \left\{(1+x).I_o(\tfrac{x}{2}) + X.I.(\tfrac{X}{2})\right\} \qquad (5)$$

$$\sigma^2_{v_o} = 2 + 2X - (\bar{v})^2 \qquad (6)$$

where I_0 and I_1 are modified Bessel functions of the first kind, of order 1 and 2, respectively.

It is clear from equation (3), that the contribution of y_n becomes negligible as the value of the signal to noise power ratio, X, becomes larger, and the performance with the linear enevelope detector tends towards that with the coherent detector. Thus the variance, $\sigma^2_{v_o}$ tends to unity, and equation (6) may be rearranged to approximate for $\bar{v}$, namely:

$$\bar{v} \approx (2X+1)^{\frac{1}{2}}$$

$$\approx \sqrt{2X} + \frac{1}{2\sqrt{2X}} \quad (X \gg 1) \qquad (7)$$

2.5 Error probability with envelope detection

The threshold value (V_o) to be applied to v_o, and the resultant equal error probability, are given by the implicit equation:

$$P_{e2} = \int_{V_o}^{\infty} \frac{d\ P(v_o)}{dv_o} . dv_o \Bigg|_{x=o} = \int_{o}^{V_o} \frac{d\ P(v_o)}{dv_o} . dv_o \Bigg|_{x=X} \tag{8}$$

The first of these functions can be evaluated directly:

$$P_{e2} = \int_{V_o}^{\infty} v_o . e^{-v_o^2/2} . dv_o$$

$$= \left[- e^{-v_o^2/2} \right]_{V_o}^{\infty}$$

$$= e^{-V_o^2/2} \tag{9}$$

Thus:

$$V_o^2 = -2 \log_e P_{e2} \tag{10}$$

and:

$$V_e^2 = -2\sigma_n^2 . \log_e P_{e2} \tag{11}$$

where V_e, is the actual threshold to be applied to v_e. Thus the threshold can be set for a required error probability when no signal is present, provided the noise level can be controlled.

Rice(ref.2) has evaluated numerically the second function of equation (8), and Barton(ref.3) plots a graph from Rice's data from which equation (8) can be evaluated for signal to noise ratio providing $(1 - P_D)$ and P_n on Bartons graph are both made equal to the assumed value of P_{e2}. This is plotted in figure 2 for values of P_{e2} between 0.3 and 10^{-4}, giving values of signal to noise ratio X, in the range 3 to 15 dB. By the use of this graph, and of equation (10), a curve of optimum threshold as a function of signal to noise ratio has been produced (figure 3).

From figures 2 and 3, it can be seen that, over the medium range of signal to noise ratios, plotted, the envelope detector requires a significantly higher optimum threshold, and gives a poorer probability of error, compared with the coherent detector. Note that the performance is slightly poorer even at very high signal to noise ratios, owing to the need to raise the threshold, because of the poorer noise rejection when the signal is not present.

Instead of choosing the threshold to equalise the two types of possible errors (i.e. equalising the areas under the tails of the distributions with, and without a signal present), Reiger(ref.4) chooses it to minimise the sum

of the two areas (i.e. to minimise the total error probability). Thus he equates to zero, the derivative with respect to V_o, of the sum of the two terms in equation (8). That is:

$$\left.\frac{d\,P(v_o)}{dv_o}\right|_{\substack{x=o\\ v_o=V_o}} = \left.\frac{d\,P(v_o)}{dv_o}\right|_{\substack{x=X\\ v_o=V_o}} \qquad (12)$$

Substitution of the relations from equation (4), and some manipulation gives:

$$X = \log_e \left(I_o \left(\sqrt{2X}\,.\,V_o\right)\right) \qquad (13)$$

This relates V_o indirectly to the signal to noise ratio, X. Evaluation shows that the values of V_o differ negligibly from the curve of figure 3, except for signal to noise ratios below 5 dB, where the difference is still small (see crosses on figure 3). This formula can be extended beyond 15 dB of signal to noise, if necessary.

Note that Reiger's figure 5 differs by 3 dB from the curve for P_{e2} in figure 2, of this report. This is because Reiger does not use true signal to noise ratio on his abscissa.

2.6 Bandpass filter

In sub-section 2.3, we showed that the sin x/x-filter-part of the matched filter, when fed with constant signal power, W, for a time, τ, produced at its output, at the end of the time, τ, *a signal component of power, W.* When fed with a noise density, N, the total noise power at the output is $\sigma_n^2 = \frac{N}{\tau}$.

If a simple bandpass filter is used in place of the matched filter, there may be two variations under the same input conditions. The signal power at the output may differ from W at time τ, due to a different transient response of the filter. The noise power may differ from $\frac{N}{\tau}$ due to a different noise band-width. If the actual signal to noise power at time, τ, is calculated (or measured) under the nominal input conditions, it will be found to be less than the matched case of $\frac{W.\tau}{N}$, and this short fall is a direct measure of the inefficiency of the bandpass filter, compared with the matched type. The curves of figures 2 and 3 are still applicable, depending on the detector type, for calculation of the required threshold and error probability. Of course the actual signal to noise ratio at the output of the filter must be used, rather than the theoretical optimum value of $\frac{W.\tau}{N}$.

Thus, for a given desired error probability, given filter type, given detector type, and given noise density, N, the required signal power at the input of the bandpass filter, can be determined, using figure 2 and the filter efficiency factor.

For an arbitrary bandpass filter, the peak signal power, at the output, may occur before (or even after) the end of the input pulse. In this case, the peak value is normally the appropriate one for determining the efficiency factor, as sampling of the derived signal is normally carried out at this time.

2.7 Efficiency of single tuned filter

To illustrate the calculation of efficiency of a filter, the single tuned filter is analysed in Appendix III. Figure 4 shows the efficiency (in decibels) plotted against $B.\tau$, the time bandwidth product. The peak efficiency of 0.815 (-0.89 dB) occurs for $B.\tau = 0.400$. This shows how insensitive the efficiency is to the actual filter shape, as the single tuned filter differs greatly from the optimum sin x/x filter, and has no side lobes at all. Obviously the side lobes are of small significance.

The fall off in efficiency for larger values of $B.\tau$ is due to the noise bandwidth (proportional to $B.\tau$) increasing faster than the signal response (which eventually levels out at unity).

For smaller values of $B.\tau$, the amplitude of the envelope at time τ is proportional to $B.\tau$. Thus the narrow filter acts as an integrator. The efficiency therefore falls off in proportion to $B.\tau$. The integration can be put to good use if the signal element is gated, prior to the filter. The system then approximates to the matched filter case, discussed in Section 2.2, and the efficiency tends towards unity, for small values of $B.\tau$. Other ways of looking at the improvement due to gating are:

(1) The gate prevents a steady state noise response from being built up. The gated noise power is considerably smaller, with a narrow filter and slow response rate. The signal is unaffected providing the gate straddles it.

(2) The Fourier spectrum of a single gate pulse is of course, a sin x/x pattern. This gives the correct filter response, and the single tuned filter acts as the integrator, as discussed above.

The gated scheme has one of the disadvantages of the original matched filter, considered. Thus the epoch of the element must be known - it cannot be derived by inspection of the response. A coherent detector is not essential, however.

2.8 Peak power and average energy

From the foregoing discussion, it is clear that the signal energy ($W.\tau$) received (and therefore transmitted) during a signal element in which a pulse is present, determines the error probability per element, and vice versa. Once the pulse length, τ, is determined, the peak power, W, is therefore also determined, independent of the number of elements transmitted per second. The average power received (and therefore transmitted) is:

$$\bar{W} = W . \tau . \frac{R_e}{R_b} . R_b . D_e \tag{14}$$

where: R_e is the number of elements transmitted per second

R_b is the number of bits of data transmitted per second

D_e is the proportion of elements containing a pulse.

For the simple coding case, where one element transmits one bit of data, R_e is equal to the bit rate (R_b) and D_e normally takes the value 0.5. This is not necessarily so, for more complex coding methods.

3. ERROR PROBABILITY WITH BI-PHASE MODULATION

With bi-phase modulation, each signal element contains a constant amplitude pulse the phase of which takes two possible values, separated by 180°. A little thought shows that the same matched filter, discussed in Section 2, is appropriate for this type of signal, but linear envelope detection cannot be used. The derived signal takes a positive or negative value, depending on the content of the element, and the optimum threshold is obviously zero, independent of noise level (a big advantage). As the difference in the mean values of the Gaussian distributions for the two element values are now doubled (assuming the same values of W, and τ) the noise amplitude must also be doubled to give the same error probability. Thus the appropriate curve for error probability versus signal to noise ratio is that for coherent detection of binary level amplitude modulation shifted to the left by 6 dB (see curve Pe_3 of figure 2).

4. INTEGRATION OF IDENTICAL ELEMENTS

4.1 General

The power, W, and element length, τ, are normally chosen to provide adequate signal to noise ratio, and error probability under the expected conditions of propagation loss and noise density. It may be, however, that for a small fraction of the operating time, conditions of noise level and/or propagation loss deteriorate. Rather than vary W or τ (which are often difficult to do in practice) it may be more acceptable to repeat the transmission of data one or more times. This will reduce the error probability, per data bit, at the expense of a reduced data transmission rate. There are many ways of processing the replicated data elements, and, as usual, the simpler the method, the less the reduction in required signal to noise ratio for a given error probability.

In this section we discuss replication processing on an element basis.

4.2 Coherent detection

If a single element is transmitted n times, one could arrange them contiguously in time. The signal to noise ratio at the output of the modified matched filter is then $\frac{W.n.\tau}{N}$, giving a reduction in signal strength of n times for the same error probability. Thus the improvement is 10 $\log_{10}$ n, decibels, and this is the maximum possible value.

If the matched filter is unmodified, we get (whether or not the elements are contiguous) n samples of the derived signal. With coherent detection, the sum of these n samples is also Gaussian distributed, of mean amplitude increased by n, and variance also increased by n. Thus the improvement in signal to noise ratio is also n, i.e. the optimum value. Thus, with coherent detection, the full improvement due to integration is obtained, i.e. there is no integration loss. The required threshold is n times that for one element.

4.3 Linear envelope detection

If the n samples of derived signal are summed in this case, the resultant distribution of noise is neither Gaussian, nor the same as that for one element (equation (4)). Marcum(ref.5) has analysed this distribution, and has calculated, numerically, the optimum threshold, and the resulting error probability. Since these are functions of the number of elements, of the signal to noise ratio before summation, and of the error probability of interest, the results are rather bulky. Marcum was examining the radar detection problem, and hence considers the two types of error separately, thus increasing the bulk further.

It has been pointed out by several authors, that the data can be condensed with reasonable accuracy, by considering integration loss as a function only of n, and the signal to noise ratio after summation (which defines the error probability). This is a very convenient set of parameters for design purposes, anyway. Figure 5 has been taken from such a set of curves presented by Kadish(ref.6). It was derived from data for a square law detector, but Marcum claims that the difference should be very small - less than 0.2 dB (figures 42 and 43 of reference 5). Figure 6 has been taken from data given by Meyer & Mayer(ref.7) and is also derived from data for a square law detector. They repeat Marcum's claim of application to the linear envelope detector. Error probability is the parameter, rather than signal to noise ratio, although there is a good correspondence via curve P_{e2} of figure 2. Note that the threshold for optimum integration is given by Marcum(ref.5, figure 57).

4.4 Post threshold integration

If the integration of n replicates takes place after the binary quantisation process, the best strategy is to take the majority decision (or choose at random, if no majority). The optimum threshold is the same as for a single replicate. The error probability with n replicates is:

$$P_{en} = \sum_{j=[\frac{n+1}{2}]}^{n} C_j^n \cdot P_e{}^j \cdot (1-P_e)^{n-j} \cdot Q_j \tag{15}$$

where $[\frac{n+1}{2}]$ is the integral part of $\frac{n+1}{2}$, and $Q_j = 0.5$, if $j=\frac{n}{2}$. $Q_j=1$, otherwise. Figure 7 shows P_{en} plotted as a function of P_e, for n = 1 to 10. Note that the results with any even, n, are the same as with the preceding odd number. Thus only an odd number of replicates would be used.

The integration loss, and total loss can be determined from figures 7 and 2. Consider the case, $P_{en} = 10^{-4}$, n = 9. For no replication, we must work at points A or B of figure 2 (for coherent and linear envelope detections, respectively). Perfect integration of 9 pulses would give an improvement of 9.55 dB in S/N, moving the operating point to C and D respectively. In fact, from figure 7, with post threshold integration, we need an input error probability of $10^{-1 \cdot 2}$, i.e. points E and F on figure 2. The integration loss is therefore the horizontal distance between points C and E, and D and F respectively. Note that the latter is larger. The total loss is the distance from the optimum point, C, to E and F. It is thus equal to the integration loss plus the detector loss (D-C) = (B-A), i.e. the detector loss at the required operational error probability (of 10^{-4}).

Figure 8 shows the post threshold integration loss, for a linear envelope detector, plotted in a way, similar to figure 6. Only odd values of n are validly represented.

5. ONE OUT OF (K + 1) CODING

5.1 General

(K + 1) elements are grouped to form a code word. They may be separated, or contiguous, but do not overlap in time. One, and only one, of the elements contains a signal pulse. The remainder contain only noise. It may be considered as a discrete pulse-position modulation system, rather than a discrete pulse-amplitude modulation system. For optimum reception, a matched filter, of the type previously considered, is fitted to each element of the code word. The one with the largest derived signal, is the one most likely to carry the signal pulse. Of course, replicates can be integrated before or after comparison of element values.

The number of bits of data, per word, is $\log_2 (K + 1)$. Hence:

$$\frac{R_e}{R_b} \cdot D_e = \frac{n \cdot (K+1)}{\log_2 (K+1)} \cdot \frac{1}{K+1} = \frac{n}{\log_2 (K+1)} \qquad (16)$$

this compares with the value, $\frac{n}{2}$ for single element coding.

5.2 Word error probability

An expression is derived in Appendix IV, for the probability of incorrect decoding of the code word, as a function of K, and the signal to noise ratio, X, at the input of a linear envelope detector, and for n = 1, i.e.:

$$P_{w1} = \sum_{j=1}^{K} \frac{(-1)^{j+1}}{j+1} \cdot C_j^k \cdot e^{-\frac{j.x}{j+1}} \qquad (17)$$

This is plotted in figure 9, for (K + 1) = 2, 4, 8 and 16.

5.3 Bit error probability

For these values (i.e. where $\log_2 (1 + K)$ is integral) we can estimate the signal to noise ratio giving a mean bit error probability, rather than a word error probability. Given a word error, each arrangement of $\log_2 (1 + K)$ bits (other than the correct one) is equally possible. Therefore the mean number of bit errors per word error is:

$$P_{b/w} = \frac{\sum_{j=1}^{\log_2 (K+1)} j \cdot C_j^{\log_2 (K+1)}}{\sum_{j=1}^{\log_2 (K+1)} C_j^{\log_2 (K+1)}} = \frac{\sum_{j=1}^{\log_2 (K+1)} j \cdot C_j^{\log_2 (K+1)}}{K} \qquad (18)$$

Evaluation shows that $P_{b/w}$ takes the values $10^{0.0}$, $10^{0.12}$, $10^{0.23}$, $10^{0.33}$ for (K + 1) = 2, 4, 8 and 16 respectively. Thus, the signal to noise ratio for a mean bit error probability of 10^{-x}, is that for a word error probability of

$10^{-(x + \log_{10} P_{b/w})}$. Note however, that there is a high correlation between the probabilities of errors of individual bits contained within one code word. For (K + 1) = 8, the probability of an error in bit number 2, is 0.5, given an error in bit number 1, and the probability of errors in both bits 2 and 3 is 0.25. This becomes very significant if the decoded bits are part of a larger code word (see Section 8).

5.4 Word error probability - decoding individual elements

Consider the case of each element being quantised by the same threshold level. The individual decisions are then combined to estimate the occupied element. If P_A is the probability a signal is missed, and P_B, the probability that a sample of noise is recorded as a signal, for any one element, then the probability of decoding the word in error is:

$$P_{w2} = P_A + (1-P_A) \left\{1 - (1-P_B)^K\right\} \tag{19}$$

$$\approx P_A + K \cdot P_B \quad (\text{if } P_A, P_B \ll 1) \tag{20}$$

The best performance is obtained if the threshold is set to adjust P_A and P_B to minimise (19). This does not make $P_A = P_B$. The problem is examined in Appendix V, for the case of linear envelope detection. The following relation is obtained:

$$X = \sqrt{2X} \cdot v_K - \log_e K - \tfrac{1}{2} \log_e 2\pi \cdot \sqrt{2X} \cdot v_K \tag{21}$$

which gives X, and v_K in terms of the parameter $\sqrt{2X} \cdot v_K$, assuming the latter is not less than 7. This proves to be the case, providing X is not less than 6 to 8 dB. Figure 10 shows v_K plotted as a function of X for K = 1, 2, 4, 8. The value for K = 1 is the same as that given in figure 3, for single element decoding (i.e. $P_A = P_B$). Appendix V determines the ratio of values for equation (20), assuming that the threshold takes the optimum value (from equation (21)) and the value which makes $P_A = P_B$. This ratio is found to be small, and approximately independent of X (over the range of 9 to 14 dB). It is well approximated by the values, $10^{-0.25}$, $10^{-0.12}$, $10^{-0.03}$ for K = 8, 4 and 2, respectively.

An examination of figure 11, which shows four possible error configurations, suggests that the error relationship of equation (19) can be improved upon. Case (a) is assumed to represent the correct code word, with one pulse detected in element 1, and the remaining K elements with no pulses detected. One can select an element at random, from those apparently eligible, after element decoding. In the cases (b) and (d) there is a finite probability of success. The probability of failure is shown at the right of figure 11. Thus the new error probability becomes:

$$P_{w3} = P_A \cdot (1-P_B)^K \cdot \frac{K}{K+1} + P_A \cdot \sum_{r=1}^{K} C_r^K \cdot P_B^r \cdot (1-P_B)^{K-r}$$

$$+ (1-P_A) \cdot \sum_{r=1}^{K} C_r^K \cdot P_B^r \, (1-P_B)^{K-r} \cdot \frac{r}{r+1}$$

$$= P_A \cdot (1-P_B)^K \cdot \frac{K}{K+1} + \sum_{r=1}^{K} C_r^K \cdot P_B^r \cdot (1-P_B)^{K-r} \cdot \left(\frac{r + P_A}{r + 1}\right) \quad (21)$$

and, taking the first term of the summation

$$\approx (1-P_B)^K \cdot \frac{K}{K+1} \cdot (P_A + \frac{K+1}{2} \cdot P_B) \quad (22)$$

Equation (21) has been plotted in figure (12), for the non-optimum case of $P_A = P_B$. A comparison of the approximations (20) and (22), shows that the same threshold correction and expected improvement factor can be applied, by substituting $\frac{K+1}{2}$ for K in figure 10.

5.5 Summary

Even with the improvement in error performance discussed in the last paragraph of the previous sub-section, the element threshold decoding is significantly inferior to the maximum amplitude decoding of sub-section 5.2. This is illustrated in Table 1, where signal to noise ratios are given for several error-word probabilities, for the several possible configurations considered. As maximum amplitude decoding has the real advantage of not requiring a threshold dependent on the signal to noise ratio, it is clearly the best choice.

Note the relatively small increase in signal to noise ratio required to transmit extra bits of information by means of the One out of (K + 1) coding scheme as K is increased. It is shown in Section 9, that, for values of K below 15, the transmitted energy per bit of information actually is reduced, as K is increased. The real penalty, of course, is in the reduction in data rate due to the rapidly increasing number of elements needed to transmit a single code word.

6. INTEGRATION LOSS WITH ONE OUT OF (K + 1) DECODING

6.1 Pre decoding integration

With pre-decoding integration of n replicates of a one out of (K + 1) code-word, the derived signal is summed, element by element. The probability distribution of the element sum, with linear envelope detection is the same one discussed in subsection 4.3. With threshold decoding of each element,

the integration loss is therefore the same as that derived in 4.3, and given in figure 6, for the required element error probability.

With maximum amplitude decoding, an analysis such as is given in Appendix IV, is necessary, using the probability distribution of the sum of n replicates. Marcum(ref.5) gives an expression for this distribution (assuming a square law detector, however) and Sarkies(ref.9, Appendix I) has used this to calculate the probability of word error with maximum amplitude decoding. Data for (K + 1) = 2 and 4, and for n = 2, 4 and 8, are shown plotted in figure 9. The integration loss has been determined and is plotted in figure 13, for (K + 1) = 4. It is almost identical to the loss for (K + 1) = 2. Again, the results for a linear envelope detector are expected to approximate closely, those for the square law detector. Note that the loss is generally smaller than for the threshold decoder.

6.2 Post decoding integration

In this case, the n replicates are separately decoded (by any method) to determine which element contains the excitation. The n results are then processed logically. If P_w is the probability of error of anyone of the n replicated words, we may calculate P_{wn}, the resultant error probability after processing, and hence determine the integration loss.

Suppose we get r errors and (n-r) correct results from the n replicates. This probability is readily calculable from the probability of individual word error. The r errors are distributed among the K empty elements. If no empty element receives (n-r) or more hits, the word is correctly decoded. If x empty elements receive (n-r) hits and none receive more, the probability of error is $\frac{x}{x+1}$. If any empty element receives more than (n-r) hits, the probability of error is unity. Thus, by evaluating these cases, and summing over all r, the total error probability is deduced. This turns out to be tedious to enumerate exactly, for n greater than 5. We will therefore determine an approximation which is also an upper limit, for these cases. The analysis is carried out in Appendix VI. The approximation is very close for small P_w, and improves as n increases.

The results are plotted in figure 14, which gives P_{wn} as a function of P_w for various values of (K + 1) and n. Note that the case, n = 2, provides no improvement at all. The integration loss is therefore 3 dB. For (K + 1) = 2, three and four replicates give the same result. The latter is rather wasteful, therefore.

The integration loss can be obtained by combining figure 14, and figure 9 (for maximum amplitude decoding) or figures 14, 12 and 2, (for threshold decoding). The former has been evaluated and the results are given in Table 2. Note that it is significantly greater than the loss for pre-detection integration, shown in figure 13. For example, the difference at $P_{wn} = 10^{-4}$, (K + 1) = 4, and n = 5 is 1.5 dB.

A somewhat surprising result is that the integration loss can be less with threshold decoding, than with maximum amplitude decoding. The overall performance is poorer, however, when the decoding loss is included. For example, if $P_{wn} = 10^{-4}$, (K + 1) = 4, the S/N required for threshold detection is 15.5 and 12.2 dB for n = 1, and n = 4. The integration loss is therefore (6.0 - (15.5 - 12.2)) = 2.7 dB, compared with 3.1 dB from Table 2. This shows the need for very careful analysis when comparing coding and detection systems.

6.3 Post decoding integration with optimum threshold detection

Two types of post decoding integration are possible with threshold detection. Firstly, each word can be decoded individually and the results processed. The relationship between word error probability and final error probability is the same as that analysed in the previous section, and illustrated in figure 14, for the maximum amplitude decoder. The resulting integration loss was also briefly discussed in the previous section. The second type of integration involves individual summation over the n replicates, of the hits in each element. The element with most hits is taken as the output decision. Where two or more elements exhibit the maximum number, the decision is based on an arbitrary selection from these elements.

The performance of the second type of process is analysed in Appendix VII, and results are given in figure 15 for various values of (K + 1) and n. Note that in this case the abscissa defines the element error probability with threshold detection, rather than the individual word error probability of figure 14. Gain and integration loss are derived as before, using figures 15 and 2. Results are given in Table 3 in the same form as for Table 2.

Note the very great reduction in integration loss, compared with the amplitude decoding case. In some cases the loss is slightly negative (-0.2 dB for $K = 1$, $n = 2$, $P_{wn} = 10^{-4}$). This appears to be genuine, and should not be regarded as exceeding a physical limitation. It represents the very efficient recovery of a lot of information discarded in the crude binary quantisation, inherent in the threshold detection process. When summed over n replicates, the effective number of quantisation levels is increased.

The recovery is so efficient that the net result can be superior to maximum amplitude detection followed by word decision integration. For example, for $n = 4$, $K + 1 = 4$, $P_{wn} = 10^{-4}$, we find that the signal to noise ratios required are: 12.2 dB, for threshold detection and word decision integration; 9.9 dB, for maximum amplitude detection and word decision integration; and 9.7 dB for threshold detection and element decision integration.

6.4 Post re-encoding integration

The output of the one-out-of-(K + 1) decoder is a decision that a particular one of the (K + 1) elements was activated. If this information is re-encoded as a normal binary code, with $m = \log_2(K + 1)$, bits, before subsequent use, it becomes feasible to carry out replication integration on a bit by bit basis, correcting each one individually. It is of interest to determine the probability of error of the integrated full binary codeword.

There are K possible error formats, each assumed to be equiprobable. In the re-encoded format, these comprise all possible arrangements of m bits other than that containing all zerors. If an error is known to exist, a little thought shows that the probability of an error in the first bit is $(K + 1)/2K$; and the probability of an error in the second, and subsequent bits, (a) given an error, and (b) given no error in the previous bit, is (a) 1/2, and (b) $(K + 1)/2(K - 1)$.

The probability of getting exactly r errors with n replicates, and exactly s, t and u errors in the first, second, and third bit of the re-encoded data is therefore approximately:

$$P(r,s,t,u) = C_r^n \cdot P_w^r \cdot (1-P_w)^{n-r} \cdot C_s^r \cdot \left(\frac{K+1}{2K}\right)^s \cdot \left(1-\frac{K+1}{2K}\right)^{r-s} \cdot$$

$$\sum_{i=0}^{s} C_i^s \cdot (\tfrac{1}{2})^s \cdot C_{t-i}^{r-s} \cdot \left(\frac{K+1}{2(K-1)}\right)^{t-i} \cdot \left(1-\frac{K+1}{2(K-1)}\right)^{r-s-t+i} \cdot$$

$$\sum_{j=0}^{t} C_j^t \cdot (\tfrac{1}{2})^t \cdot C_{u-j}^{r-t} \left(\frac{K+1}{2(K-1)}\right)^{u-j} \cdot \left(1-\frac{K+1}{2(K-1)}\right)^{r-t-u+j} \qquad (23)$$

where: $C_o^o = 1$

The probability of getting an error after replication integration is:

$$P_{wn} = \sum_{r=\left[\frac{n+1}{2}\right]}^{n} \cdot \sum_{s=0}^{r} \cdot \sum_{t=0}^{r} \cdot \sum_{u=0}^{r} \cdot P(r,s,t,u) \cdot Q(s,t,u) \qquad (24)$$

where:

$Q = 0$; if s and t and u $< {}^n/_2$

$Q = 1$; if s or t or u $> {}^n/_2$

$Q = 1-(\tfrac{1}{2})^x$; if x of s,t,u $= {}^n/_2$

Equation (24) has been evaluated for n = 2, 3, 4, 5, 7, 10 and (K + 1) = 2, 4, 8, and for a range of values of P_w. Curves are not given, since in all cases, the performance is either equal or inferior to that for replication integration of the (K + 1) elements, determined in the previous sub-section. If (K + 1) = 2, the performance is the same, as would be expected since only one bit is involved. If n = 3, the performances are almost identical. For n = 4, and larger n, the performance is significantly inferior. Even values of n are slightly inferior to the preceding odd value.

There seems to point, then, in using replication integration after re-encoding except for large values of (K + 1), where the reduction in complexity may be significant. It is the confounding of information during the re-encoding which results in a loss of integration performance.

7. GOLAY CYCLICAL CODE

7.1 General

We have considered several ways of transmitting data with binary level, signal elements over a Gaussian noisy channel, using more than the minimum required number of elements, to obtain reasonable error probability, with reduced signal power. These include repetition of data words, and grouping of elements into (K + 1) element codewords, with only one active element.

There are more efficient ways of allocating redundant elements, one of which is the use of cyclic codes, capable of correcting a certain number of element errors. In this report, only the Golay cyclic code is considered, as it is a moderately powerful typical example without being too complex for convenient analysis or practical application. Cyclic codes, and the Golay code in particular have been examined theoretically and practically by this Author in references 10 and 11. This report draws heavily on the theory given in reference 10.

The Golay code consists of a block of 12 incoming data elements to which are added 11 redundant or parity elements, to form a codeword of 23 elements. The decoder generates a syndrome (an 11 bit word) from the received 23 element block. Error correction relies on the fact that every distinct error word of less than weight 4 (i.e. four ones in the 23 elements) superimposed on the transmitted word, produces a distinct syndrome. Correction merely involves subtraction (or addition, modulo 2) of the known error word corresponding with the syndrome. Error words of weight 4 or more, produce syndromes non distinct from those for lower weight errors. Therefore these errors are erroneously corrected, and the error is not identified.

7.2 No replication of Golay code

If the 23 transmitted elements are statistically independent (that is, not grouped in any way, such as in a (K + 1) codeword), the probability of a decoding error is just the probability of receiving four or more errors in 23 elements, using threshold detection. Thus:

$$P_{G1} = \sum_{j=4}^{23} C_j^{23} \cdot P_e^{\,j} \cdot (1-P_e)^{23-j} \tag{25}$$

This is shown plotted in figure 16 (curve n=1) as a function of the element error probability, P_e. Note that P_{G1} is the probability of the whole block of 12 data bits being in error. The distribution of the errors among the 12 bits is not considered here.

7.3 Pre decoding integration

If the replicates are integrated, element by element, before, or immediately after element threshold detection, the element error probabilities are still independent and equation (25) still applies. The integration loss is as derived in sub-sections 4.3 and 4.4, respectively.

7.4 Post decoding integration

If the replicates are integrated logically, after Golay decoding, we need to know the probability of getting each syndrome, since identical syndromes give the same correction, but not necessarily the same erroneous result for replicated codewords. As more than one error word, of differing inherent

probability, (if errors of weight greater than 3 are included) can produce the same syndrome, this has a very significant effect on the probability of getting more than one identical erroneous decoded word, and hence on the probability of error in the logical integration process.

Suppose the replicated transmitted word is:

$$C = G \times B$$

where B is the 12 element data word, and G, the Golay generation function.

Suppose one of the received words is:

$$C + E = G\,(B + F) + S$$

where E is the 23 element error word, and S is the syndrome, generated by the decoder. F is determined by the equation. (S is the residue when C + E is divided by G).

The correction device determines the assumed error, determined by S, as:

$$E' = G \times F' + S$$

where E' is of weight less than 4. Adding this, modulo 2, to the received word gives:

$$C + E + E' = G\,(B + F + F')$$

and the output of the Golay decoder is $(B + F + F')$.

The problem reduces to the determination of the probability distribution of $(F + F')$, or $(E + E')$, as a function of the number of elements in error in the received code (i.e. the weight of E). E' has a weight less than 4 since it is chosen to correct an error, E, of weight less than 4. It is shown in Appendix VIII, that the probability of an error word, E, of weight, r, correcting to a particular error word $(E + E')$ of weight, w, is:

$$X_{w,r}/C_r^{23}$$

and an expression for evaluation of $X_{w,r}$ is given. Since $(E + E')$ is a valid cyclic codeword (it leaves no residue on division by G), its weight, w, can take only certain values. For the Golay code, these are (with number, N_w of such codewords in brackets): 0(1), 7(253), 8(506), 11(1288), 12(1288), 15(506), 16(253), 23(1). Note that the values of N_w are derived by inpsection of the entire set of codewords. The weight, r, can take any value between 0 and 23, depending only on the number of elements in error. For many of these, however, $X_{w,r}$ is zero. Table 4 lists the values of $X_{w,r}$.

The probability of getting any weight, w, corrected error word, with random error in each element of the codeword, is therefore:

$$P_w = \sum_{r=0}^{23} C_r^{23} \cdot P_e{}^{r} \cdot (1-P_e)^{23-r} \cdot \frac{N_w \cdot X_{w,r}}{C_r^{23}} \qquad (26)$$

where the C_r^{23} terms cancel. All words of weight, w, are equally likely. Values of $\log_{10} P_w$ are given in Table 5.

The probability of getting x_0, x_7, x_8, x_{11}, x_{12}, x_{15}, x_{16}, x_{23} corrected words of weight 0, 7, 8, ... 23, with n replicates is:

$$P' = P_o^{x_0} \cdot P_7^{x_7} \ldots\ldots P_{23}^{x_{23}} \cdot C_{x_0}^{n} \cdot C_{x_7}^{n-x_0} \cdot C_{x_8}^{n-x_0-x_7} \ldots\ldots C_{x_{23}}^{x_{23}} \qquad (27)$$

where

$$\sum x_i = n$$

Of course there are only 8 non zero values of x_i, for the Golay code.

Given x_w corrected words of weight, w, the marginal probabilities of getting exactly x_o, and more than x_o identical error words of weight, w, can be calculated, as in Appendix VI. As these have to be combined for all seven values of x_w, $w > 0$, in relation (27), and the results again combined with values of P', for all possible distributions of the x parameters, the evaluation of the resulting probability of error is tedious even to enumerate, let alone calculate, if n is larger than 5.

The evaluation can be simplified by taking only the more significant terms of equation (27), but this results in poor accuracy at the levels of final probability of most interest, e.g. from 10^{-2} to 10^{-4}. Table 6 enumerates the exact results for values of n between 2 and 5, from which the following relations are derivable. These are exact for n = 2 and 3, and very good approximations for n = 4 and 5. In the latter, we ignore the contribution in excess of fraction (n - 1)/n, for $x_o = 1$.

$$\left.\begin{aligned}
P_{G2} &= \sum_{w=7}^{23} P_w = Z && ;\ n = 2\\
&= Z^2(2-Z) + (1-Z)\sum_{w=7}^{23}\frac{P_w^2}{N_w} && ;\ n = 3\\
&= Z^3\ (3-2Z) + 3(1-Z)^2\sum_{w=7}^{23}\frac{P_w^2}{N_w} && ;\ n = 4\\
&= Z^4\ (4-3Z) + 15(1-Z)^2\sum_{w=7}^{23}\frac{P_w^2}{N_w} && ;\ n = 5
\end{aligned}\right\} \qquad (28)$$

These are shown plotted in figure 16. Note that the results for n = 2 are identical with those for n = 1, (no replication).

Integration loss for envelope detection, derived from figures 16 and 2, is given in Table 7. Note that post decoding integration is significantly inferior to pre decoding integration.

7.5 Post decoding integration using correction weight

The strategy for selecting the chosen codeword from the n replicates, as used in the previous sub-section, involves decoding and correcting each replicate, individually, followed by selection of the one appearing most often. Where more than one word shares the maximum number, the selection is made randomly from this sub-group. Some better way than chance, for selecting from multiple choices, would give better error performance, particularly in the case of $n = 2$, where chance gives no improvement over $n = 1$.

A better way exists, for a cyclic code. It involves the inspection of the weight of the correction word, E', used with each of the multiple choice results, and the selection of that choice with the lowest weight. The improvement in performance comes about as follows.

Consider the Golay code, for two replicates. The multiple choice arises when one replicate is decoded correctly and the other incorrectly. The correct result comes from 0, 1, 2 or 3 bit errors ($r = 0, 1, 2, 3$), giving 0, 1, 2 or 3 bit weighting for the correction, E. The zero weight is, of course, more probable than the others at the signal to noise ratios of interest. The most probable source of the incorrect word is from 4 bit errors ($r = 4$) and this always (for the Golay code) gives a weight 3 correction. A 5-bit error gives weight 2 and 3 corrections. Table 8 enlarges on this arrangement (up to $r = 9$), which follows from an examination of cases 1 to 4 of Appendix VIII. Also shown, is the probability of getting an error of each weight, given r bit errors in the received word. Chance is still used for selection between words of equal weights.

For each probability of getting a bit error, the total probability of getting a weight 0, 1, 2 or 3 correction with and without a correct result (P_0, P_1, P_2, P_3, Q_0, Q_1, Q_2, Q_3) can be evaluated (the Q values by summation over $r = 4$ to 23 bit errors, although values of r beyond 8 have little significance).

A little thought shows that the total decoding error probability, for $n = 2, 3$, is given by:

$$P_{Gn} = Q_0 (P_0 + 2P_1 + 2P_2 + 2P_3) + Q_1 (P_1 + 2P_2 + 2P_3) + Q_2 (P_2 + 2P_3) + Q_3 . P_3 + (Q_0 + Q_1 + Q_2 + Q_3)^2 \quad ; n = 2 \qquad (29)$$

$$P_{Gn} = 3Q_0 (Q_1 + Q_2 + Q_3)(P_0 + 2P_1 + 2P_2 + 2P_3) + 3Q_1 (Q_2 + Q_3)(P_1 + 2P_2 + 2P_3) + 3Q_2 . Q_3 (P_2 + 2P_3) + Q_0^2 (2P_0 + 3P_1 + 3P_2 + 3P_3) + Q_1^2 (2P_1 + 3P_2 + 3P_3) + Q_2^2 (2P_2 + 3P_3) + 2 Q_3^2 . P_3 + (Q_0 + Q_1 + Q_2 + Q_3)^3 \quad ; n = 3 \qquad (30)$$

These equations have been evaluated, and are shown dotted, in figure 16. The improvement in performance is quite significant, particularly for $n = 2$.

There is an error mode, for $n = 3$ or greater, which has been ignored. For $n = 3$, this occurs when one replicate is correct, and two are incorrect, and the latter two give identical correction words. The latter is automatically chosen. Its probability of occurrence is, from Table 6:

$$P'_{Gn} = 3 \cdot P_0 \cdot \sum_{w} \frac{P_w^2}{N_w} \qquad (31)$$

where P_O, P_W and N_W have the same significance as in the previous subsection. The true total error probability for $n = 3$, is given to a good approximation, by the summation of equations (30) and (31). The difference is negligible for total probabilities greater than about 10^{-5}. Below this point, the curve (of figure 16) falls off at a slope at 8 decades/decade, rather than the value of 11, resulting from equation (30) alone.

It is clear that complete evaluation for $n \geq 4$ is likely to be extremely tedious, although approximations could be obtained for the lower range of probabilities. The significance of the improvement obtainable when the effect of equation (31) is included, is not obvious. The limited number of examples in figure 16, suggest that the improvement over the main range of interest, is somewhat less than that obtained by increasing the value of n by one, without using the correction weight criterion. Crude argument in favour of this theory is as follows.

The probability of getting zero correct decoded replicates must be very low for a satisfactory communication channel. If greater than n/2 replicates decode correctly, there is no possibility of error under either criterion of selection. For $1 \leq s \leq n/2$, the number of incorrectly decoded replicates $(n - s)$ equals or exceeds the number of correctly decoded replicates, s. For an error to occur, however, at least s of the $(n - s)$ must give rise to the same (incorrectly decoded) output word. As the probability of getting a particular correction is low, most errors occur through exactly s of the $(n - s)$ giving the same correction. Thus, the new criterion resolves most of the most probable errors. If an extra replicate were taken (without using the new criterion), the probability is high that it would give a correctly decoded result. If so, it would also resolve the same set of most probable errors. Hence there is some justification for expecting the new criterion to give roughly the same improvement in total error probability, as for an increase of one in the number of replicates.

8. CASCADED (K + 1) ELEMENT, AND GOLAY CODES

8.1 No interleaving of bits

If the 23 bits of a Golay codeword are transmitted in the form of 23, 12, 8 or 6 words of a 2, 4, 8 or 16 element, one out of (K + 1) code, the result depends on how many of the bits of a (K + 1) codeword are associated with the same Golay codeword. This follows from the fact that the bit errors in a (K + 1) codeword are not independent (see Section 5.3). With no interleaving (all bits associated with one Golay word) the performance is readily calculated. We calculate the probability of getting 1, 2, 3, 4 etc. return errors out of the 23, 12, 8 and 6 needed to provide enough bits for a 23 bit Golay word (with one over, in the last three cases). The marginal probability of getting 4 or more bits in error in the Golay word is then calculated, allowing for the extra bit which cannot affect the Golay word. This is tedious, but straightforward, and the result is given in Table 9.

The overall performance is shown in figure 17, where Golay decoder error probability is plotted against (K + 1) codeword error probability for (K + 1) equal to 2, 4, 8 and 16. The first of these is the same relation as for n = 1, in figure 16, and this follows from the fact that, for (K + 1) = 2, only one bit is transmitted per (K + 1) codeword.

At first sight, the curve for (K + 1) = 16 seems to give little improvement for the extra bits transmitted, being only marginally better than the line of unit slope, passing through the origin. In fact, of course, the output refers to a block of 12 data bits, whereas the input refers to only 4 bits. If the curves are combined with those of figure 9, to determine the signal to noise ratio required for maximum amplitude decoding, we find that 8.9, 10.7, 11.4 and 13.1 dB are required for (K + 1) = 2, 4, 8 and 16, for a Golay error of 10^{-4}. Where peak power is a constraint, (K + 1) = 2 is clearly the best choice. When the energy per Golay word is calculated, taking account of the total number of activated elements, we find relative results of +2.2, +1.0, 0.0, +0.4 dB for the four cases. Thus (K + 1) = 8 gives minimum energy.

Much better performance can be achieved by interleaving the bits, as described in the next sub-section.

8.2 Full interleaving of bits

If 23 words of a (K + 1) code each contribute one bit to $\log_2(K + 1)$ separate 23-bit Golay codewords, the effect of the correlation between bit errors in the (K + 1) codeword are greatly reduced. The mean decoded Golay word error probability is given by:

$$P_{G3} = \sum_{r=4}^{23} C_r^{23} \cdot P_w^{\,r} \cdot (1-P_w)^{23-r} \cdot \sum_{s=4}^{r} C_s^r \cdot P_b^s \cdot (1-P_b)^{r-s} \cdot Y(r,s) \qquad (32)$$

where P_b (2/3, 4/7, 8/15 for (K + 1) = 4, 8, 16) is the probability that the bit allocated to one particular Golay word is in error, given that the (K + 1) codeword is in error; and Y(r,s) = 1. P_b is obtained by listing the possible outcomes of bits, all being of equal probability. Curves of P_{G3} versus P_w are shown in figure 18. Note the very small difference for (K + 1) equal to 4, 8 and 16. If (K + 1) equals 2, only one bit is carried, and the interleaving process is meaning-less. In general,

$$P_b = \frac{K + 1}{2K}$$

Because of the correlation in the error of the bits of the (K + 1) codeword, there is a correlation between the errors in the $\log_2(K + 1)$ separate Golay codewords. This means that there is a much greater probability of getting multiple Golay word errors than would be expected from the mean probability of an error in any one word. It does not mean, unfortunately, that the probability of getting at least one error in the $\log_2(K + 1)$ words is significantly reduced, as might at first sight be expected, from the increased tendency for errors to cluster. The tendency is still small enough to have negligible effect, as is shown below.

Equation (32), with Y(r,s) = 1, gives the probability of getting 4 or more errors in the bits allocated to any particular one of the Golay words (and hence of getting an error in that decoded word). The probability of getting 4 or more errors in another particular Golay word, given an error in the

first, is therefore obtained by substituting for Y(r,s) in equation (27):

$$Y(r,s) = \sum_{t=4}^{r} \sum_{j=0}^{s} C_j^s \, (P_a)^j . (1-P_a)^{s-j} . C_{t-j}^{r-s} . (P_c)^{t-j} . (1-P_c)^{r-s-t+j} . Y_1(r,t) \quad (33)$$

where $Y_1(r,t)$ is unity, P_a is the probability of getting an error in the second bit of a (K + 1) codeword, given an error in the first, and P_c is the probability of getting an error in the second bit, given no error in the first (but given an error in at least one bit of the codeword). P_a and P_c are evaluated by listing all (equi probable) bit error combinations (ignoring the null error one). Thus $P_a = \frac{1}{2}$, and $P_c = 1, 2/3, 4/7$ for (K +1) = 4, 8, 16. Similarly, the probability of getting 4 or more errors in the third Golay word, given errors in the previous two, is obtained by substituting in equations (32) and (33):

$$Y_1(r,t) = \sum_{u=4}^{r} . \sum_{i=0}^{t} C_i^t . (\tfrac{1}{2})^t . C_{u-i}^{r-t} . (P_d)^{u-i} . (1-P_d)^{r-t-u+i} . Y_2(r,u) \quad (34)$$

where again, $Y_2(r,u)$ is unity. $P_d = 2/3, 4/7$ for (K + 1) = 8, 16.

This process can be continued indefinitely, for large (K + 1) codes, but soon becomes too time consuming for accurate evaluation even by digital computer. Results have been obtained for (K + 1) = 4 and 8. These have been rearranged to give the probability of one only, two only, and three only, Golay word errors, out of the two or three words in the group. Figure 18 shows curves plotted from the results. Some points worth noting are:

(a) The curves for only one error are almost identical for (K + 1) = 4, 8. (This is a chance effect, not to be extrapolated for higher K, see below).

(b) The curves for mean error probability are almost identical for (K + 1) = 8, 16, and not much different for (K + 1) = 4. (This is not a chance effect; see below).

(c) The curves for more than one error are well below the mean (by approximately a factor of 10). They are nevertheless much closer than for no correlation.

(d) The curves for only one error have a maximum point, at high error probability. This is due to the probability of multiple errors exceeding that for one only.

Approximate values for the mean, and one-word-only error probabilities can be obtained as follows. Rearranging equation (32), with Y = 1; and with $P_b = \frac{K+1}{2K}$:

$$\bar{P} = \sum_{r=4}^{23} C_r^{23} . P_w{}^r . (1-P_w)^{23-r} . \frac{1}{2^r} . (1-\tfrac{1}{K})^r . \sum_{s=4}^{r} C_s^r . (1 + \frac{2}{K-1})^s \quad (35)$$

This clearly tends to a limit as K tends to infinity. This has been evaluated and plotted in figure 18.

The mean error probability:

$$\bar{P} = \frac{P_1 + 2P_2 + 3P_3 + \ldots\ mP_m}{m}$$

where m is the number of Golay words (i.e. $\log_2(K + 1)$), and P_i is the probability of getting exactly i words in error. Therefore, by rearrangement:

$$P_1 = m.\bar{P}.\left(1 - \frac{2P_2 + 3P_3 + \ldots + mP_m}{m.\bar{P}}\right)$$

$$\approx m\bar{P} \qquad (36)$$

since $P_i/\bar{P} \ll 1$, if $i \geqslant 2$ (figure 18).

The probability of getting at least one word in error, is:

$$P_{\geqslant 1} = P_1 + P_2 + \ldots + P_m$$

$$\approx P_1$$

$$\approx m.\bar{P} \qquad (37)$$

This is virtually the same as for the uncorrelated case, where:

$$P_{\geqslant 1} \approx P_1 \approx 1 - (1-\bar{P})^m \approx m.\bar{P}$$

8.3 Replication of cascaded codes

If the integration of the replicates takes place before the assembly of the Golay codeword, and its subsequent decoding, the integration loss is that applicable to the stage at which integration takes place. All such stages have been considered in previous sections.

In Section 7.4, we have considered the integration performance for the post-Golay-decoding stage, in terms of the probability of an element error (assumed to be independent of other element errors). Figure 16 shows the results for 1 to 5 replicates. With the cascaded coding system, the Golay elements are completely independent, providing full interleaving is used, because each bit is derived from a different (K + 1) codeword. (There is, of course, some slight dependence between Golay codewords from the same set of (K + 1) codewords). Thus, for the fully interleaved case, the results of Section 7.4 and figure 16 are applicable, providing the following relationship between the (K + 1) decoding error probability, P_w, and the equivalent element error probability, P_e, is used.

$$P_e = \frac{K + 1}{2K} . P_w \qquad (38)$$

$(K + 1)/2K$ is the proportion of $(K + 1)$ codeword errors which give an error in a particular element of the Golay codeword.

If full interleaving is not used, there is some dependence between the error probabilities for different elements of one Golay word. This is tedious to evaluate, and is not considered in this report. The error in assuming independence is not likely to be very large, and the poor performance of the non-interleaved case makes it unlikely to be used in practice.

8.4 Replication within the Golay group

Where the $\log_2 (K + 1)$ bits from one codeword are spread between the same number of Golay words (full interleaving), data word replicates could be allocated to some or all of these words. This results in a considerable loss in performance, compared with replication outside the Golay group, as is illustrated in the following examples. Within-group replication should therefore be avoided.

Consider the case $(K + 1) = 4$, with two bits of data and a group of two Golay words. If these words are known to contain the same data, and the bits are allocated to corresponding positions in each word, then both bits of a pair are known to be identical, and only 2 elements of the $(K + 1)$ code are active. The optimum decoding must then assume that $(K + 1) = 2$, and the two Golay words become identical. Therefore only one need be decoded. From figures 18, 17 and 9, the signal to noise ratio for a final Golay word error probability of 10^{-4}, is reduced from 9.4 dB to 8.8 dB (assuming maximum amplitude decoding). This is a poor return for halving the data rate.

If the bits are allocated to random positions in the two Golay words, we must use normal decoding of the codewords, with $(K + 1) = 4$. There is a significant correlation between the number of bits in error in the two words, however. Since we have only two replicates, we must use the correction weight method of integration, discussed in sub-section 7.5, if any improvement is to be obtained. The performance is analysed, approximately, as follows.

The probability of getting r errors in 23, $(K + 1)$ codewords, and of getting s and t bit errors in the first and second Golay replicates, is:

$$\begin{aligned} P(r,s,t) &= C_r^{23}.(P_w)^r.(1-P_w)^{23-r}.C_s^r.(2/3)^s.(1/3)^{r-s}.C_{t-r+s}^s.(\tfrac{1}{2})^{t-r+s}.(\tfrac{1}{2})^{r-t} \\ &= C_r^{23}.(P_w)^r.(1-P_w)^{23-r}.(1/3)^r.C_s^r.C_{t-r+s}^s \end{aligned} \qquad (39)$$

The probability of getting a decoding error using the correction weight method is:

$$P_G = \sum_{r=4}^{23} \sum_{s=0}^{r} \sum_{t=0}^{r} P(r,s,t) \,.\, Q(s,t) \qquad (40)$$

where Q is derived from Table 8 and is given in Table 10, for $s < t$. If the values of s and t are interchanged, Q is unchanged.
Note that Q is actually non-zero for the penultimate line of Table 10, so the results calculated from this table are slightly optimistic. These are illustrated in figure 19.

From figures 19 and 9 it follows that the required signal to noise ratio for a final error probability of 10^{-4}, is 9.2 dB, (compared with 9.4 dB with no replication), again using maximum amplitude decoding. This is an even poorer return for halving the data rate.

The deleterious effect of the error correlation with within-group replication, described above, becomes easy to understand by examining only one error contribution from Table 10, namely the last one. The probability that s and t are both greater than 3 is, of course, the probability of getting both words in error, in the non-replicated case, with full interleaving. This is illustrated by the curve (4,2) of figure 18. Comparison with figure 19, shows that this term is quite significant, and it is certainly a limiting effect on the improvement with replication. As discussed in sub-section 8.2, this probability is very much higher than if the two words were not derived from the same group of (K + 1) codewords.

9. COMPARISON OF CODING SCHEMES

9.1 Criteria for comparison

Whereas direct binary threshold detection produces a single bit of data from a single code element, and each bit is statistically independent of any other (as far as errors are concerned), the (K + 1) and Golay coding systems produce a group of statistically dependent bits from a group of elements. There is, therefore, no universal ground of comparison. The performance will depend on how the bits are grouped into words, to suit the particular application.

In many applications, the bits are grouped into words such that an error in any single bit makes the whole word more or less useless. This applies to words representing arithmetic quantities in digital computation, or navigational co-ordinates. Words of 4, 8, 12, 16, 24 and 32 bits are often used in these applications. Since the Golay code naturally produces 12 bits of data, and the other schemes, considered in this report, can also produce 12 bit words by grouping integral multiples of the natural groups, a 12-bit data word has been chosen for comparison purposes. Longer words can be made up from the 12-bit ones, as desired, making use of the fact that the 12-bit word errors are statistically independent.

Since the element configuration has been fixed at rectangular pulses of fixed duration and amplitude, containing a large number of cycles of constant frequency carrier, and since the output has been defined as a 12-bit data word with a fixed, required, probability that one or more bits are in error, and since the competition is restricted to Gaussian noise, added in the linear part of the transmission system, the bases for comparison reduce to the number of elements required per data word (code rate), the signal power transmitted within an active element, and the proportion of active elements (the duty cycle). In most applications, of course, it is desirable that the peak transmitted power, the time required to transmit a data word, and the energy consumed by the transmitter, per data word, should all be minimised. The relation between these parameters has already been discussed in Section 2.8.

For a given path loss, noise level, and quasi-matched filter, the signal power required to be transmitted is proportional to the signal to noise ratio required at the output of the filter (assuming this point is ahead of any non linear device, such as a square-law or linear envelope detector).

The type of carrier detector (coherent, linear envelope, or square law) affects the distribution of signal plus noise at its output, and therefore affects the performance (at a given signal to noise ratio) of the element detectors (binary threshold, or maximum amplitude). For comparison purposes, a linear envelope detector is assumed in this section.

9.2 Data word of 12 independent bits

Appendix IX examines the case of a data word formed by m elements, of independent error probability. The words are replicated n times and the most numerous result accepted (equally numerous results being resolved by random selection). Some approximations are made to simplify the analysis, but these should have a negligible effect on the result. Figure 20 shows the results plotted for m = 12, and n = 1,2,3,4,5,6,7 and 9. Note that two replicates give no improvement in error performance, and that an odd number of replicates tends to be slightly more efficient than an even number, in using the information available.

9.3 Data word from (K + 1) words

If $12/\log_2(K + 1)$ codewords of the one out of (K + 1) variety are combined to form a 12 bit data word, there is a dependence between the errors in bits common to a sub word. An analysis of the n-replicate situation could be carried out as for sub-section 9.2 and Appendix IX, but instead of having 12 types of error words of equal error probability (the words with 1 to 12 bits in error) we find there are 31, 35 and 35 for (K + 1) equal to 4, 8 and 16, respectively. (For (K + 1) equal to 2, there is only one bit per sub-word, and the performance is the same as for sub-section 9.2, providing the subword error probability is substituted for the element probability. In fact, from figure 12, these are identical for (K + 1) = 2).

For one replicate, the evaluation of error probability is simple, being given by:

$$P_D = 1 - (1 - P_W)^{12/\log_2(K + 1)} \qquad (41)$$

This is shown plotted in figure 21. Note, again, that the curve for (K + 1) = 2, is the same as that (for n = 1) in figure 20. The curves for other values of (K + 1) have the same slope, but are displaced slightly.

A rough idea of the improvement due to replication, could perhaps be obtained by applying the vertical offset between the n-lines in figure 20, to the appropriate (K + 1) line (for n = 1) in figure 21.

9.4 Candidate coding schemes

9.4.1 Figure 22 shows the possible coding schemes derivable from the types discussed in this report. The input consists of sequential derived signals (Section 2), in analogue or multi level digital form, appropriate to the sequence of received code elements. The output consists of 12-bit data words. The switches define alternate data paths, depending on the coding scheme. All two-section boxes include the possibility of some form of replication integration.

9.4.2 The simplest scheme involves only boxes 1, 2, 3, 4 and 17. The binary threshold detector determines whether each element is active or not, under the control of the threshold generator, and produces one bit per element. Twelve bits are combined to form a data word. Replication integration is possible on an element (or bit) basis, before and after binary threshold detection (boxes 1, and 2), and on a word basis, after data word assembly (box 4).

9.4.3 The second scheme uses binary threshold detection with (K + 1) word coding, and includes boxes 7, 8, 9 and 10 between 2 and 3. Replication integration is also possible before (K + 1) decoding (box 8) and on a word, or bit basis, after (K + 1) decoding (boxes 9 and 10). Box 8 differs from box 2, in that all (K + 1) counts are presented simultaneously to the decision logic, to produce one decision, compared with (K + 1) individual decisions for box 2.

9.4.4 The third scheme uses maximum amplitude detection, instead of binary threshold detection, with (K + 1) word coding, and involves boxes 5, 6, 9, 10, 3 and 4. Replication integration in box 6, is a true integration of continuous variables, rather than a summation of two level decisions. (Replication integration in box 1, is likewise).

9.4.5 The fourth scheme uses Golay coding, but not (K + 1) coding. It therefore necessarily uses binary threshold detection, and consists of boxes 1, 2, 11, 12, 13, 14, 15 and 16. Additional replication integration is possible after Golay decoding (box 16), with or without the use of error correction weight to resolve multiple maximum counts, or on a word basis before Golay decoding (box 12). The latter is clearly inferior in performance to both block 2 and block 16 integration, since it must show a greater error count, and is therefore not considered further. It has not been analysed in this report.

9.4.6 The fifth scheme uses (K + 1) and Golay coding in cascade, and incorporates binary threshold detection. It is comprised of boxes 1, 2, 7, 8, 9, 10, 11, 12, 13, 14, 15 and 16. No extra replication integrations are involved.

9.4.7 The sixth scheme replaces the binary threshold detection with maximum amplitude detection, and is comprised of boxes 5, 6, 9, 10, 11, 12, 13, 14, 15 and 16.

9.4.8 Table 11 relates the replication-decision boxes of figure 22, with the graphs giving the error performance, and the section in which it is derived and discussed.

9.5 Comparison of non-replicated schemes

There are six possible schemes considered in the previous sub-section, all using linear envelope detection, and each with 3 to 7 possible methods of integrating replications. Those four that use (K + 1) coding, also have four values of (K + 1) (i.e. 2, 4, 8, 16) to be considered. Replication integration is a straight trade off between data rate and signal to noise ratio, with energy consumption per data word increasing proportional to the number of replications. As such, it is less efficient than more complex types of redundancy coding, and should be considered mainly as an adaptive measure, easily introduced to handle the occasional situation of abnormally poor signal to noise ratio. We first, therefore, consider non-replicated schemes.

(K + 1) coding is mainly a trade off between data rate and energy per data word, although signal to noise ratio is reduced somewhat (about 2.5 dB) if maximum amplitude decoding is used. It has the big advantage that no threshold need be generated.

The eighteen cases, without replication, are listed in Table 12. Full interleaving is assumed for the cascaded (K + 1) and Golay schemes. An error probability of 10^{-4} for the whole data word is assumed, and the required signal to noise ratio, within an active element, is calculated from the graphs specified in the table. The code rate (data bits per element), and the transmitted energy for the 12-bit data word (relative to the non-redundant scheme, number 1) are also presented.

It is clear, from the table, that scheme 6 is best, both for minimum signal to noise ratio, and for minimum energy. The best signal to noise ratio, of 8.9 dB, occurs for (K + 1) = 2. The least energy requirement occurs for (K + 1) = 16, and would be reduced for even larger values of (K + 1), although the code rate suffers significantly, and the signal to noise ratio rises. Assuming the energy requirement is less important than signal to noise ratio, and code rate, (K + 1) = 4 would seem to be the best choice. This involves a penalty of 0.5 dB in S/N, while preserving the same data rate, and gaining 2.5 dB in energy saving, compared with (K + 1) = 2.

Case 6, with (K + 1) = 4, is therefore chosen for further examination (with replication). Figure 23 shows 12-bit error probability plotted versus S/N, for this scheme. Note that, from Table 12, we have traded a reduction in code rate of 4, for a S/N improvement of 7.2 dB, and an energy reduction of 3.7 dB. If the same code rate reduction were achieved by replication (n = 4), we would expect a S/N improvement of 4.7 dB (6 dB less the integration loss from figure 5), and an energy increase of 6.0 dB, assuming the most efficient integration process (box 1, of figure 22) is used. The advantages of sophisticated coding methods are very great!

9.6 Replication of standard coding scheme

The standard coding scheme (optimised in the previous sub-section, for no replication) is now considered with replication. From figure 22, the only possibilities for replication integration, are boxes 6, 9, 10, 12 and 16. Boxes 10 and 12 have previously been rejected (sub-sections 6.4 and 9.4.5) as being inefficient. Only boxes 6, 9 and 16 will therefore be considered here.

Figure 23 shows curves of 12-bit word error probability plotted against signal to noise ratio, for values of n in the range 1 to 4. These curves have been derived from figures 18 and 9, for box 6; figures 18, 14 and 9, for box 9; and figures 16 (modified by equation (38)) and 9. The curve for n = 4, and box 16, is estimated, assuming the correction weighted result is slightly poorer than the non-weighted result for n = 5, in figure 16. The curves for n = 2, box 6, and n = 3, box 9, are close enough to be combined in figure 23. The curves for n = 2, box 9, and n = 1 are, of course, identical.

As expected, from the rule of thumb that the less non-linear processing the better the integration performance (or the less the integration loss), box 6 gives better performance than box 9, which gives better than box 16. The only exception is for n = 2, where box 16 is better than box 9.

9.7 Replication of other coding schemes

Other schemes can be analysed in the same general manner, using information available in this report.

10. CONCLUSION

The characteristics of data transmission by means of binary level elements of a constant frequency carrier have been derived from first principles. Matched filtering is shown to give optimum detection of the element levels, in the presence of Gaussian noise. Various means of matched and quasi-matched filtering are discussed, together with various types of carrier detector and element decision devices. The theory is adequate for estimation of the performance losses due to non-ideal matching and to non-coherent carrier detection, in practical applications.

The important characteristics of a data link are the data rate (in data bits per second), the probability of error (per bit, or group of bits), the peak power required to be transmitted, and the total energy transmitted per bit. By using various encoding schemes, it is possible to trade off one characteristic against another, but usually with less than perfect efficiency. This report has considered in detail, methods of trading reduced data rate for reduced peak power, namely by repeating data several times (replication), by grouping elements so that only one out of (K + 1) contain a carrier pulse, by grouping data bits with redundant bits to form an error correctable cyclic (Golay) code, and all combinations of these schemes. Results are given to permit evaluation of any particular method. Table 12 summarises the performance of all combinations not involving replication, assuming a 12-bit word error probability of 10^{-4}.

The (K + 1) code essentially trades data rate for reduced energy per bit, with little effect on peak power (or signal to noise ratio) as K increases. It has the very great advantage, however, of permitting maximum amplitude decoding of the elements, rather than threshold decoding. This avoids the need of generating a threshold (dependent on the signal to noise ratio) and also gives a direct improvement of about 2.5 dB in peak power (almost independent of K).

The Golay code produces an improvement of 4.9 dB in peak power, and 2.1 dB in energy per bit, for a reduction of 2 in data rate. The same reduction due to replication can produce an improvement of 0 to 3 dB in peak power, with an increase of 3 dB in energy per bit. Replication is therefore very inefficient, compared with more complex coding methods, and should not be used for a basic link. It can be useful, however, for adaptive control of the link characteristics.

The concatenation of (K + 1) and Golay codes gives an improvement of 6.7 dB in peak power, and 3.7 dB in energy per bit (with K = 3) with a reduction in data rate of 4. The effect of various forms of replication are given in figure 23, when applied to this concatenated code. With integration of 2 and 4 replicates, prior to maximum amplitude detection, the signal to noise ratio can fall to as low as 7.2 and 5.1 dB respectively, for an error probability of 10^{-4} per 12-bit word, compared with 9.4 dB for the non-replicated case. Of course, the energy consumption rises by 3 and 6 dB respectively.

Intuition tends to suggest that replication integration should give a large improvement in peak power performance, when carried out after a complex decoding process, such as the Golay decoder, because of the already low error rate. Analysis shows this to be false. The more non-linear processes of carrier detection, element detection, and decoding are carried out before replication integration, the poorer the signal to noise ratio recovery. It would seem not unreasonable to extrapolate this conclusion to other similar coding schemes, not considered in this report.

In general, then, a good working rule (subject, of course, to detailed analysis in a particular case) would seem to be as follows. Use the longest error correcting code that can reasonably be decoded (such as the Golay code) in cascade with the (K + 1) code (for ease of element detection) and with maximum amplitude decoding. Determine the signal to noise ratio needed for the required error probability. Design the element length, quasi-matched filter, and peak power to give the necessary signal to noise ratio. Use replication (preferably with integration carried out as near as possible to the linear part of the receiver) only for adaptive control in situations where it is acceptable to trade a reduced data rate for an abnormal reduction in signal to noise ratio. For post decoding integration, and also when cascading codes, ensure that error samples are not statistically dependent.

NOTATION

B	data word vector
B	bandwidth of filter
C	Golay code vector
C^n_j	combinations of n, j at a time
De	proportion of elements containing a pulse
E	Golay actual error word
E′	Golay assumed error word (used for correction)
F	Golay data word error
F′	Golay data word after correction
G	Golay generating vector
K	number of empty elements in an orthogonal group
N	spectral power density
N_w	number of Golay words of weight w
P	general probability variable
$\bar{P}$	mean word error probability
P_A	probability that a hit is missed
P_B	probability of noise registering a hit
P_D	probability of detection
P_G	probability of Golay word being in error
Pb	see Pb/w
Pb/w	probability of a particular bit being in error, given a word is in error
Pe	probability of an element being decoded in error
Pn	probability of false detection
Pw	probability of a (K + 1) codeword being in error
Pwn	Pw after integration of n replicates
Q	probability of error

Q	error weighting function
Rb	bits per second
Re	elements per second
S	Golay syndrome
S/N	signal to noise power ratio
V_e	element threshold level
V_o	normalised V_e
W	pulse power
X	signal to noise power ratio
X_{wr}	probability of an error word of weight, r, giving a correction word of weight, w
Y	probability function
Z	see equation (28)
j	dummy variable
j	imaginary operator
m	number of bits in (K + 1) codeword ($\log_2(K + 1)$)
n	number of replicates
n	number of delay line taps
r	dummy variable
r	number of words actually in error
r	weight of error word
s, t, u	number of errors occurring in bits 1, 2 and 3
t	time epoch of element
v_e	element signal envelope
v_o	normalised, v_e
$\bar{v}_o$	mean of v_o
v_K	see Appendix V

w	weight of error correction word
$\left.\begin{array}{l} x \\ y \end{array}\right\}$	samples from a Gaussian distribution of zero mean
$\sum$	summation symbol
σ	standard deviation, r.m.s. noise voltage
τ	element duration

REFERENCES

No.	Author	Title
1	North, D.O.	"An Analysis of the Factors which Determine Signal/Noise Discrimination in Pulsed-Carrier Systems". Proc. IEEE, $\overline{51}$, p 596, April 1963
2	Rice, S.O.	"Mathematical Analysis of Random Noise". Bell System Technical Journal, $\overline{23}$, p 282, July 1944
3	Barton, D.K.	"Radar System Analysis". Prentice Hall, 1964, p 17
4	Reiger, S.	"Error Probabilities of Binary Data Transmission Systems in the Presence of Random Noise". IRE National Convention Record, pt.8 p 72, 1953
5	Marcum, J.I.	"A Statistical Theory of Target Detection by Pulsed Radar". Trans. of IRE. IT-6, No.2, April 1960
6	Kadish, J.E.	"Pulse Radar Range". Selenia Technical Memorandum, 65/001 Rev.1, January 1967
7	Meyer, D.P. and Mayer, H.A.	"Radar Target Detection. Handbook of Theory and Practice". Academic Press, 1973
8	Spiegel, M.R.	"Mathematical Handbook of Formulas and Tables". McGraw Hill, 1968
9	Sarkies, K.W.	"Coding for Communication Over a Totally Incoherent Non Fading Channel with Peak and Average Power Restrictions". WRE-TN-1728 (WR&D), November 1976
10	Hayward, J.	"The Generation and Correction of Cyclic Codewords Using the Theory of Polynomial Describing Functions". WRE-TN-1482 (WR&D), September 1975
11	Hayward, J.	"Burst and Random Error Correction for the Golay Code Using Shared Registers". WRE-TN-1626 (WR&D), June 1976

APPENDIX I

MATCHED FILTER RESPONSE TO RECTANGULAR PULSE IN WHITE NOISE

Consider a rectangular pulse of length, τ, carrier angular frequency, ω_c, and mean power, W. Then (at a point of unity impedance):

$$v = \sqrt{2W} \,.\, \cos\,(\omega_c . t + \phi) \tag{I.1}$$

The matched filter can consist of a multiplication process, with a reference wave, followed by an integration process, over the duration, τ.

$$v_o = \frac{\sqrt{2}}{\tau} . \int_t^{t+\tau} v \,.\, \cos\,(\omega_r . t) . dt \tag{I.2}$$

Thus:

$$v_o = \frac{\sqrt{2}}{\tau} . \int_t^{t+\tau} \sqrt{2W} \,.\, \cos\,(\omega_c t + \phi) \,.\, \cos \omega_r t \,.\, dt$$

$$= \frac{\sqrt{W}}{\tau} . \int_t^{t+\tau} \{\cos\,(\overline{\omega_c - \omega_r} \,.\, t + \phi) + \cos\,(\overline{\omega_c + \omega_r} \,.\, t + \phi)\}$$

$$= \frac{\sqrt{W}}{\tau} . \left[\frac{\sin\,(\overline{\omega_c - \omega_r} \,.\, t + \phi)}{\omega_c - \omega_r} + \frac{\sin\,(\overline{\omega_c + \omega_r} \,.\, t + \phi)}{\omega_c + \omega_r} \right]_t^{t+\tau}$$

$$= \frac{\sqrt{W}}{\tau} . \left\{ \frac{2.\cos(\overline{\omega_c - \omega_r} \,.\, (t + {}^{\tau}/2) + \phi) \,.\, \sin \overline{\omega_c - \omega_r} \,.\, {}^{\tau}/2}{\omega_c - \omega_r} \right.$$

$$\left. + \frac{2 \,.\, \cos(\overline{\omega_c + \omega_r} \,.\, (t + {}^{\tau}/2) + \phi) \,.\, \sin \overline{\omega_c + \omega_r} \,.\, {}^{\tau}/2}{\omega_c + \omega_r} \right\}$$

$\therefore$

$$\frac{v_o}{\sqrt{W}} = \cos\,(\overline{\omega_c - \omega_r} \,.\, (t + {}^{\tau}/2) + \phi \,.\, \frac{\sin \overline{\omega_c - \omega_r} \,.\, {}^{\tau}/2}{\overline{\omega_c - \omega_r} \,.\, {}^{\tau}/2}$$

$$+ \cos\,(\overline{\omega_c + \omega_r} \,.\, (t + {}^{\tau}/2) + \phi \,.\, \frac{\sin \overline{\omega_c + \omega_r} \,.\, {}^{\tau}/2}{\overline{\omega_c + \omega_r} \,.\, {}^{\tau}/2} \tag{I.3}$$

In the perfectly matched case, $\omega_r = \omega_c$, $\phi = 0$, and:

$$\frac{v_o}{\sqrt{W}} = 1 + \frac{\cos 2\omega_c (t + \tau/2) \cdot \sin \omega_c \tau}{\omega_c \tau}$$

$\therefore$

$$1 - \frac{1}{\omega_c \tau} \leqslant \frac{v_o}{\sqrt{W}} \leqslant 1 + \frac{1}{\omega_c \tau}$$

Thus, for $\omega_c \tau \gg 1$, we may neglect the second term of equation (I.3).

Consider the response to white noise, at a density of N watts per Hz of bandwidth (all noise is assumed to be allocated to positive Hz). In a small bandwidth, Δf, we have an equivalent signal of: $N.\Delta f$.
Averaged over a large number of samples (of pulses) the equivalent phase angle, ϕ, is uniformly distributed. Noting that the average value of $\cos^2\phi$ is ½, we find that the total variance in the output, v_o, due to white noise, is, from equation (I.3):

$$\sigma^2_{v_o} \approx \int_0^\infty \frac{N.df}{2} \cdot \frac{\sin^2 \overline{\omega - \omega_r} \cdot \tau/2}{(\omega - \omega_r)^2 \cdot \tau^2/4}$$

Putting:

$$f = \omega/2\pi$$

$$x = \overline{\omega - \omega_r} \cdot \tau/2$$

and assuming $\omega_r \gg 0$

$$\sigma_{v_o}^2 \approx \int_{-\infty}^{\infty} \frac{N}{2} \cdot \frac{\sin^2 x}{x^2} \cdot \frac{1}{\pi . \tau} \cdot dx$$

But Spiegel (reference 8, relation 15.36) gives:

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} \cdot dx = \pi$$

Therefore:

$$\sigma^2_{v_o} \approx \frac{N}{2\tau} \qquad \text{(I.4)}$$

Since the correlation process (equation (I.2)) is linear, the output noise has the same distribution as the input noise, i.e. Gaussian.
Note that the output noise has a magnitude equivalent to the output of a filter of noise-bandwidth:

$$B = \frac{1}{2\tau} \tag{I.5}$$

fed with the same white noise of energy density, N watts per Hz.

Note also that the signal to noise ratio at the output of the matched filter is:

$$X_o = \frac{W}{\sigma^2_{v_o}} = 2 \, . \, \frac{W.\tau}{N} \tag{I.6}$$

where $W.\tau$ is the energy received during the signal element of length, τ, when the signal is actually present.

The Gaussian probability density is given by the well known relation:

$$\frac{d\ P(x)}{dx} = \frac{1}{\sigma.\sqrt{2\pi}} \, . \, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{I.7}$$

where μ is the mean value, and σ^2 the variance of the distribution.

APPENDIX II

MULTIPLE TAPPED DELAY LINE AS A MATCHED FILTER

Consider a rectangular pulse of length, τ, carrier angular frequency, ω_c, and mean power, W, fed into a lossless delay line, of length, τ, with n taps, equally spaced. Suppose the signals at the taps are fed to a summing network. The output of the network is a carrier wave with amplitude steps occurring at intervals of τ/n. If τ/n is a multiple of the period of the carrier wave, the result is an increasing staircase, which roughly resembles an integration process, in terms of the amplitude of the carrier. We may evaluate the amplitude of the wave, at the end of the input pulse, without this assumption.

$$v_s(t+\tau) = \frac{1}{n}\sum_{i=0}^{n-1}\sqrt{2W}.\cos(\omega_c(t+\tau - i.\frac{\tau}{n})+\phi)$$

$$= \frac{\sqrt{2W}}{2n}.\sum_{i=0}^{n-1}\left(e^{j(\omega_c(t+\tau-i.\frac{\tau}{n})+\phi} + e^{-j(\omega_c(t+\tau-i.\frac{\tau}{n})+\phi)}\right)$$

$$= \frac{\sqrt{2W}}{2n}.e^{j(\omega_c(t+\tau)+\phi)}.\sum_{i=0}^{n-1}e^{\frac{-j.i.\tau.\omega_c}{n}} + \frac{\sqrt{2W}}{2n}.e^{-j(\omega_c(t+\tau)+\phi)}.\sum_{i=0}^{n-1}e^{\frac{j.i.\tau.\omega_c}{n}}$$

$$= \frac{\sqrt{2W}}{2n}.e^{j(\omega_c(t+\tau)+\phi)}.\frac{1-e^{-j.\tau.\omega_c}}{1-e^{\frac{-j.\tau.\omega_c}{n}}} + \frac{\sqrt{2W}}{2n}.e^{-j(\omega_c(t+\tau)+\phi)}.\frac{1-e^{j.\tau.\omega_c}}{1-e^{\frac{-j.\tau.\omega_c}{n}}}$$

$$= \frac{\sqrt{2W}}{2n}.\frac{e^{\frac{j.\tau.\omega_c}{2}}-e^{\frac{-j.\tau.\omega_c}{2}}}{e^{\frac{j.\tau.\omega_c}{2n}}-e^{\frac{-j.\tau.\omega_c}{2n}}}.\left\{e^{j(\omega_c(t+\tau-\frac{\tau}{2}+\frac{\tau}{2n})+\phi)} + e^{-j(\omega_c(t+\tau-\frac{\tau}{2}+\frac{\tau}{2n})+\phi)}\right\}$$

$$= \frac{\sqrt{2W}}{n}.\cos(\omega_c(t+\frac{n+1}{2n}.\tau)+\phi).\frac{\sin\frac{\tau.\omega_c}{2}}{\sin\frac{\tau.\omega_c}{2n}} \qquad \text{(II.1)}$$

Putting

$$\omega_r = \frac{2\pi.m.n}{\tau}$$

$$= \frac{\sqrt{2W}}{n} . \cos(\omega_c(t + \frac{n+1}{2n} . \tau)+\phi).(-1)^{m(n-1)} . \frac{\sin\overline{\omega_c-\omega_r}.\tau/2}{\sin\overline{\omega_c-\omega_r}.\frac{\tau}{2n}}$$

(II.2)

$$= \sqrt{2W} . \cos(\omega_c(t + \frac{n+1}{2n}\tau)+\phi).(-1)^{m(n-1)} . \frac{\sin\overline{\omega_c-\omega_r}.\tau/2}{\overline{\omega_c-\omega_r}.\tau/2} . \frac{\overline{\omega_c-\omega_r}.\frac{\tau}{2n}}{\sin\overline{\omega_c-\omega_r}.\frac{\tau}{2n}}$$

This is, of course, a signal at the carrier frequency. Its amplitude contains the same sin x/x term as the main term of equation (I.3). Providing n is large enough, the last term is approximately unity for signals (including noise) in the vicinity of ω_r.

The carrier frequency signal can be "detected" using a multiplication process with a reference signal, as for Appendix I, but without the integration. The instantaneous output at $(t+\tau)$, assuming that a low pass filter removes the double carrier frequency component, is approximately:

$$v_o(t+\tau) = \sqrt{2} . v_s(t+\tau) . \cos\omega_r t$$

$$\approx \sqrt{W} . \cos(\overline{\omega_c-\omega_r}.t + \frac{n+1}{2n}.\omega_c.\tau+\phi).(-1)^{m(n-1)} . \frac{\sin\overline{\omega_c-\omega_r}.\tau/2}{\overline{\omega_c-\omega_r}.\tau/2} \quad \text{(II.3)}$$

It is clear from equations (II.3) and (I.3) that the multiple tapped delay line followed by summation and coherent detection gives virtually the same matched filter performance as coherent detection followed by integration. In both cases, the reference signal must be identical in frequency, and correctly phased. Again the process is linear, and the output noise has a Gaussian distribution.

APPENDIX III

EFFICIENCY FACTOR OF SINGLE TUNED FILTER

Consider a sinewave voltage source, feeding a lossless, parallel L-C circuit via a resistance, R. The steady state response is given by:

$$\frac{v_o}{v_i} = \frac{1}{1 + R \cdot \dfrac{j\omega L + \dfrac{1}{j\omega C}}{j\omega L \cdot \dfrac{1}{j\omega C}}} \qquad \text{(III.1)}$$

$$= \frac{1}{1 + j \cdot \dfrac{R}{\omega L} \cdot (\omega^2 LC - 1)}$$

$$= \frac{1}{1 + 2j \cdot \dfrac{R}{\omega_0^2 . L} \cdot (\omega - \omega_0)\left(1 + \dfrac{\omega-\omega_0}{2\omega_0}\right)\left(1 + \dfrac{\omega-\omega_0}{\omega_0}\right)^{-1}} \quad ; \quad \omega_0^2 = \frac{1}{LC}$$

If $R \gg \omega_0 L$ (i.e. if the circuit Q is large) $\left|\dfrac{v_o}{v_i}\right|$ is small unless $\dfrac{\omega-\omega_0}{\omega_0}$ is also small. Thus the last two bracketted terms may be approximated by unity. Substituting:

$$B = \frac{\omega_0^2 . L}{2\pi R}$$

(i.e. the 3 dB bandwidth in Hz)

$$\frac{v_o}{v_i} \approx \frac{1}{1 + j \cdot \dfrac{\omega - \omega_0}{\pi . B}} \qquad \text{(III.2)}$$

The equivalent noise bandwidth is given by:

$$\frac{1}{2\pi}\int_0^\infty \left|\frac{v_o}{v_i}\right|^2 . \, d\,\omega$$

$$\approx \frac{1}{2\pi}\int_0^\infty \frac{d\,\omega}{1 + \dfrac{(\omega-\omega_0)^2}{\pi^2 . B^2}}$$

$$\approx \frac{B}{2} \cdot \int_{-\frac{\omega_0}{\pi . B}}^{\infty} \frac{dx}{1+x^2}$$

$$\approx \frac{B}{2}\left[\arctan x\right]^{\infty}_{-\frac{\omega_o}{\pi . B}}$$

$$\approx \frac{\pi . B}{2} \text{ (assuming } \frac{\omega_o}{\pi . B} \text{ is very large)} \qquad \text{(III.3)}$$

The La Place transform of the transfer function is obtained by substituting s for $j\omega$ in equation (III.1).

$$\frac{\bar{v}_o}{\bar{v}_i} = \frac{1}{1 + RC\ (s + \frac{1}{SLC})}$$

Putting: $v_i = \sin \omega_o . t \quad ; \quad 0 \leqslant t \leqslant \tau \quad ;$ otherwise zero:

$$\bar{v}_i = \frac{\omega_o}{s^2 + \omega_o^2}\ (1 - e^{-s\tau})$$

$$\bar{v}_o = \frac{\omega_o}{s^2 + \omega_o^2} \cdot (1 - e^{-s\tau})\left(1 - \frac{s^2 + \omega_o^2}{s^2 + s/RC + \omega_o^2}\right)$$

Taking the inverse La Place transform, and noting that $R.C = \frac{1}{2\pi B}$:

$$v_o = \sin\omega_o t - \frac{e^{-\pi . B . t}}{\sqrt{1 - \frac{\pi^2 . B^2}{\omega_o^2}}} \cdot \sin\omega_o t \cdot \sqrt{1 - \frac{\pi^2 B^2}{\omega_o^2}} \qquad \text{(III.4)}$$

$$\text{for } 0 \leqslant t \leqslant \tau$$

At $t = \tau$, the phase shift between the two sinewaves is approximately:

$$\omega_o . \tau \cdot \frac{\pi^2 . B^2}{2\omega_o^2} = (B.\tau) \cdot \frac{2\pi B}{\omega_o} \cdot \frac{\pi}{4} \ll \frac{\pi}{4}$$

Thus the response at time, τ, is approximately:

$$v_o = (1 - e^{-\pi . B . \tau}) \cdot \sin\omega_o t \qquad \text{(III.5)}$$

Compared with the matched filter response of unity, and noise bandwidth $\frac{1}{\tau}$, the efficiency of the single tuned filter is:

$$\eta = \frac{2}{\pi . B . \tau} \cdot (1 - e^{-\pi . B . \tau}) \qquad \text{(III.6)}$$

APPENDIX IV

MAXIMUM AMPLITUDE DECODING OF ONE OUT OF (K + 1) ELEMENT CODE

Assuming linear envelope detection, the probability that the K empty elements have their derived signals below the level, V_o, is:

$$\left(\int_o^{V_o} v.e^{-v^2/2}.dv\right)^K = \left(1-e^{-v_o{}^2/2}\right)^K = \sum_{j=0}^{K} C_j^K.\left(-e^{-v_o{}^2/2}\right)^j$$

Therefore, the probability that noise in at least one empty element, exceeds the level obtained in the occupied element, is from equation (4):

$$P_w = \int_o^{\infty} V_o.e^{-\frac{v_o{}^2+2X}{2}}.I_o(\sqrt{2X}.V_o).\left(1-\sum_{j=0}^{K} C_j^K.(-e^{-V_o^2/2})^j\right).dV_o$$

Substituting $y = 2X.V_o{}^2$:

$$P_w = \frac{e^{-X}}{4x}.\sum_{j=1}^{K} C_j^K.(-1)^{j+1}.\int_o^{\infty} I_o(\sqrt{y}).e^{-\frac{(j+1).y}{4x}}.dy$$

Spiegel(ref.8, relations 24.94 and 24.32) gives:

$$\int_o^{\infty} e^{-ax}.J_o(b.\sqrt{x}).dx = (1/a).e^{-b^2/4a}$$

$$I_o(x) = J_o(j.x) \quad (j^2 = -1)$$

$\therefore$

$$P_w = \frac{e^{-X}}{4x}.\sum_{j=1}^{K} C_j^K.(-1)^{j+1}.\frac{4x}{j+1}.e^{\frac{X}{j+1}}$$

$$= \sum_{j=1}^{K} \frac{(-1)^{j+1}}{j+1}.C_j^K.e^{-\frac{jx}{j+1}} \qquad \text{(IV.1)}$$

APPENDIX V

OPTIMUM ERROR PROBABILITY FOR ONE OUT OF (K + 1) CODE (THRESHOLD DETECTION)

Consider the code word of (K + 1) elements, only one of which contains a signal. If P_A is the probability a signal is missed, and P_B is the probability that an erroneous signal is recorded, for one element, then the total probability of obtaining a decoding error is:

$$\begin{aligned} P &= P_A + (1-P_B).(1 - (1-P_B)^K) \\ &\approx P_A + K.P_B \end{aligned} \tag{V.1}$$

There is no question of equalising mark and space errors in this code, so the optimum threshold gives minimum total error probability, i.e.:

$$\frac{dP_K}{dv_K} \approx \frac{dP_{AK}}{dv_K} + K \cdot \frac{dP_{BK}}{dv_K} = 0 \tag{V.2}$$

where v_K is the threshold used to decode the elements.

Assuming linear envelope detection:

$$\frac{dP_A}{dv} = v \cdot e^{-\frac{v^2+2X}{2}} \cdot I_o(\sqrt{2X}.v) \tag{V.3}$$

But:

$$I_o(\sqrt{2X}.v) \approx \frac{e^{\sqrt{2X}.v}}{\sqrt{2\pi}.\sqrt{2X}.v} \text{ if } \sqrt{2X}.v \geqslant 7$$

(Spiegel,(ref.8, relation 24.107)

Some manipulation gives:

$$\frac{dP_A}{dv} \approx \sqrt{\frac{v}{2\pi\sqrt{2X}}} \cdot e^{-\frac{(v-\sqrt{2X})^2}{2}} \quad ; \quad \sqrt{2X}.v \geqslant 7 \tag{V.4}$$

Then:

$$P_{AK} = \int_{-\infty}^{v_K} \frac{dP_A}{dv}.dv = \int_{-\infty}^{v_1} \frac{dP_A}{dv}.dv + \int_{v_1}^{v_K} \frac{dP_A}{dv}.dv \tag{V.5}$$

$$P_{BK} = \int_{v_K}^{\infty} \frac{dP_B}{dv} \,.\, dv = \int_{v_K}^{\infty} v \,.\, e^{-v^2/2} \,.\, dv = e^{-v_K^2/2} \qquad (V.6)$$

Substituting from equations (V.5), (V.6), (V.4), (V.3) into equation (V.2) to obtain the optimum threshold, :

$$\sqrt{\frac{v_K}{2\pi\sqrt{2x}}} \,.\, e^{\frac{-(v-\sqrt{2x})^2}{2}} - K.v_K.e^{-v_K^2/2} = 0$$

whence, by manipulation:

$$X = \sqrt{2}x.v_K - \log_e K - \tfrac{1}{2}\log_e 2\pi.\sqrt{2x}.v_K \qquad (V.7)$$

This gives X in terms of the parameter ($\sqrt{2}x.v_K \geqslant 7$). We then can obtain v_K, $\sqrt{\frac{v_K}{\sqrt{2x}}}$, and ($\sqrt{2}x-v_K$), also in terms of the same parameter. Figure 10 shows these quantities plotted against X, for K = 1, 2, 4, 8. The condition for valid approximation (equation (V.4)) is met for X greater than 6 dB to 8 dB, which covers cases of main interest.

To determine the significance of the error reduction at threshold, v_K, compared with v_1, we determine the relation:

$$\frac{P_K}{P_1} = \frac{P_{AK} + K.P_{BK}}{P_{A1} + K.B_{B1}}$$

Substituting values from equations (V.5) and (V.6), and noting that (from Section 2.5):

$$\frac{P_K}{P_1} = \frac{e^{-v_{1/2}^2} + \int_{v_1}^{v_K} \frac{dP_A}{dv} \,.\, dv + K.e^{-v_{K/2}^2}}{(1+K).e^{-v_{1/2}^2}}$$

$$> \frac{1+e^{v_{1/2}^2}.\left\{K.e^{-v_{K/2}^2} + \frac{\sqrt{\frac{v_1}{\sqrt{2X}}} + \sqrt{\frac{v_K}{\sqrt{2X}}}}{2}.\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{2X}-v_1} .e^{-v^2/2}.dv - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{2X}-v_K} e^{-v^2/2}.dv\right]\right\}}{1+K} \qquad (V.8)$$

Note that the mean value of $\sqrt{\frac{v}{\sqrt{2X}}}$ is used, since this changes quite slowly with K (figure 10) and its removal from the integral leaves the integral error function which is well tabulated (e.g. Spiegel, page 257). Equation (V.8) is evaluated at fixed values of X, and v_1, v_K, $\sqrt{\frac{v_1}{\sqrt{2X}}}$, $\sqrt{\frac{v_K}{\sqrt{2X}}}$, $\sqrt{2X}-v_1$, $\sqrt{2X}-v_K$ are all taken from the graph (figure 10). Over the range of X of 9 to 14 dB (the range of main interest):

$$\log_{10} \frac{P_K}{P_1} \approx -0.25, -0.12, -0.03 \text{ for } K = 8, 4 \text{ and } 2, \text{ respectively} \qquad (V.9)$$

This improvement is relatively small.

APPENDIX VI

POST DETECTOR INTEGRATION OF ONE OUT OF (K + 1) CODE

Suppose a decoder, of a (1 + K) element codeword, has an individual error probability of P_w. Suppose the post decoding logic gives a marginal final error probability of $P(n,r)$, given r decoding errors from n replicates. Then the full final error probability is:

$$P_{wn} = \sum_{r=o}^{n} C^n_r \cdot P_w{}^r \cdot (1-P_w)^{n-r} \cdot P(n,r) \qquad \text{(VI.1)}$$

The probability that s of the r errors occur in one particular element, is:

$$C^r_s \cdot \left(\frac{1}{K}\right)^s \cdot \left(\frac{K-1}{K}\right)^{r-s}$$

If $S > r/2$, there is insufficient errors to provide s or more errors in any other element. Therefore the probability of getting exactly one element with s errors is:

$$K \cdot C^r_s \left(\frac{1}{K}\right)^s \cdot \left(\frac{K-1}{K}\right)^{r-s} \qquad \text{(VI.2)}$$

In this case,

$$P(n,r) = \sum_{s=n-r+1}^{r} K.C^r_s.\left(\frac{1}{K}\right)^s.\left(\frac{K-1}{K}\right)^{r-s} + 0.5\ K.C^r_{n-r}.\left(\frac{1}{K}\right)^{n-r}.\left(\frac{K-1}{K}\right)^{2r-n} \qquad \text{(VI.3)}$$

providing $n-r > r/2$; i.e. $r < \frac{2n}{3}$

$$P(n,r) = 0 \ ; \text{ if } r < n-r \ ; \text{ i.e. } r < n/2$$

$P(n,r)$ is tedious to evaluate if $r \geqslant \frac{2n}{3}$, apart from the two values:

$$P(n,n) = 1 \qquad \text{(VI.4)}$$

$P(n,n-1) = 1$; if $K < n-1$ (since one element must contain two or more errors)

$$P(n,n-1) = 1 - 1 \cdot \frac{K-1}{K} \cdot \frac{K-2}{K} \cdots\cdots \frac{K-(n-2)}{K} \cdot \frac{1}{n} \ ; \text{ otherwise.} \qquad \text{(VI.5)}$$

(examining the condition that no element has more than one error).

The following table shows the situation for various n.

n	$\left[\frac{n+1}{2}\right]$	$\left[\frac{2n-1}{3}\right]$	$(n-1)$	n
2	1	1	1	2
3	2	1	2	3
4	2	2	3	4
5	3	3	4	5
6	3	3	5	6
7	4	4	6	7
8	4	5	7	8
10	5	6	9	10

Thus, relations (VI.3), (VI.4) and (VI.5) can be used to give exact values of P_{wn} for $2 \leqslant n \leqslant 5$. For higher values of n, an upper bound is obtained by using (VI.3) and the approximate value, $P(n,r) = 1$, for $\left[\frac{2n+2}{3}\right] \leqslant r \leqslant n$. The results are plotted in figure 14, for various values of $(K + 1,n)$.

APPENDIX VII

POST DETECTOR INTEGRATION, (K + 1) CODE, THRESHOLD DETECTION

Consider n replicates of a (K + 1) element code word, decoded with a threshold detector, and threshold crossings summed element by element. Suppose P_A is the probability that a signal is missed, and P_B, the probability that an erroneous signal is recorded.

The probability of getting exactly r hits in the active element, is:

$$R_r = C_r^n \cdot (1-P_A)^r \cdot P_A^{n-r} \qquad \text{(VII.1)}$$

The probability of getting exactly r hits in any one inactive element, is:

$$S_r = C_r^n \cdot P_B{}^r \cdot (1-P_B)^{n-r} \qquad \text{(VII.2)}$$

The probability of getting at least (r + 1) hits in any one inactive element, is:

$$\sum_{j=r+1}^{n} S_j$$

The probability of getting at least one inactive element with (r + 1) or more hits, is:

$$T_r = \sum_{i=1}^{K} C_i^K \cdot \left(\sum_{j=r+1}^{n} S_j\right)^i \cdot \left(1- \sum_{j=r+1}^{n} S_j\right)^{K-i}$$

$$= 1 - \left(\sum_{j=0}^{r} S_j\right)^K \qquad \text{(VII.3)}$$

The probability of getting exactly m inactive elements, each with exactly r hits; and (K-m) elements, each with less than r hits, is:

$$Q(r,m) = C_m^K \, (S_r)^m \cdot \left(\sum_{j=0}^{r-1} S_j\right)^{K-m} \qquad \text{(VII.4)}$$

The total probability of decoding n replicates in error, is therefore:

$$P_{wn} = \sum_{r=0}^{n} R_r \cdot \left\{ T_r + \sum_{m=1}^{K} \frac{m}{m+1} \cdot Q(r,m) \right\}$$

$$= R_o \left\{ 1-(1-P_B)^{n.K} + \frac{K}{K+1} \cdot (1-P_B)^{n.K} \right\} + \sum_{r=1}^{n} R_r \cdot \left\{ 1 - \left(\sum_{j=0}^{r} S_j \right)^K \right.$$

$$\left. - \sum_{m=1}^{K} \frac{1}{m+1} \cdot C_m^K (S_r)^m \cdot \left(\sum_{j=0}^{r-1} S_j \right)^{K-m} + \left(\sum_{j=0}^{r} S_j \right)^K - \left(\sum_{j=0}^{r-1} S_j \right)^K \right\}$$

$$= R_o \cdot \left\{ 1 - \frac{1}{K+1} \cdot (1-P_B)^{n.K} \right\} + \sum_{r=1}^{n} R_r \cdot \left\{ 1 - \left(\sum_{j=0}^{r-1} S_j \right)^K \right.$$

$$\left. - \sum_{m=1}^{K} \frac{1}{m+1} \cdot C_m^K (S_r)^m \cdot \left(\sum_{j=0}^{r-1} S_j \right)^{K-m} \right\} \qquad \text{(VII.5)}$$

Substituting $BIN(K_1, K_2, n, P) = \sum_{j=K_1}^{K_2} C_j^n \cdot P^j \cdot (1-P)^{n-j}$

$$P_{wn} = BIN(0,0,n,1-P_A) \cdot \left(1 - \frac{1}{K+1} \cdot (1-P_B)^{n.K}\right)$$

$$+ \sum_{r=1}^{n} BIN(r,r,n,1-P_A) \cdot \left\{ 1-(1-BIN(r,n,n,P_B))^K \right. \qquad \text{(VII.6)}$$

$$\left. - \sum_{m=1}^{K} \frac{1}{m+1} \cdot C_m^K \cdot BIN(r,r,n,P_B)^m \cdot (1-BIN(r,n,n,P_B))^{K-m} \right\}$$

If P_A and P_B are very small, we find the assymptote for P_{wn} from (VII.3), (VII.4) and (VII.5), namely:

$$P_{wn} \to \frac{K}{K+1} \cdot R_o + \sum_{r=1}^{n} R_r \cdot \left\{ K.S_{r+1} + 0.5\ K.S_r \right\}$$

$$\to \frac{K}{K+1} \cdot P_A{}^n + \frac{K}{2} \cdot \sum_{r=1}^{n} C_r^n\ P_A{}^{n-r} \cdot C_r^n\ P_B{}^r$$

$$\to \frac{K}{2} \cdot \left\{ \sum_{r=0}^{n} (C_r^n)^2 \cdot P_A{}^{n-r} \cdot P_B{}^r - \frac{K-1}{K+1} \cdot P_A{}^n \right\} \qquad \text{(VII.7)}$$

Neglecting the second term of equation (VII.7), which is much smaller than the first, if $P_A \approx P_B$, and $n > 1$, (VII.7) is symmetrical in P_A and P_B, and hence its minimum value occurs for $P_A = P_B$. Under this assumption:

$$P_{wn} \to \frac{K}{2} \cdot P_A{}^n \cdot \left\{ \sum_{r=0}^{n} (C_r^n)^2 - \frac{K-1}{K+1} \right\} \qquad \text{(VII.8)}$$

APPENDIX VIII

PROBABILITY OF AN ERROR WORD PRODUCING A GIVEN GOLAY - CORRECTED ERROR WORD

We consider the inverse case. Consider a particular corrected error word, $(E + E')$ of weight, w, and number of elements, m. Suppose this is to be transformed into an original error word, E, of weight, r, by the addition (modulo 2) of the coset leader, E', of weight less than half the minimum distance of the cyclic code, d. (For the Golay code, m = 23, d = 7).

$E + E'$ contains w Ones and m-w Zeroes

E contains r Ones and m-r Zeroes

E' contains $0 \leqslant x < \frac{d}{2}$ Ones and m-x Zeroes

Case 1 $(0 < w < m;\ \ r \geqslant w)$

$(r - w + i)$ Ones from E' cover $(r - w + i)$ out of $(m - w)$ Zeroes of $(E + E')$

i Ones from E' cover i out of w Ones of $(E + E')$

to give r Ones in E.

This can be done in

$$X_{w,r} = \sum_i C^{m-w}_{r-w+i} \cdot C^{w}_{i} \text{ ways;} \tag{VIII.1}$$

where i is chosen to satisfy:

$$0 \leqslant r - w + i \leqslant m-w$$
$$0 \leqslant i \leqslant w$$
$$r - w + 2i < \frac{d}{2}$$

i.e.

$$\left.\begin{aligned} i &\geqslant 0 \\ i &\leqslant w \\ i &\leqslant m-r \\ i &< \frac{\frac{d}{2} - (r-w)}{2} \end{aligned}\right\} \tag{VIII.2}$$

Case 2 ($0 < w < m$; $w \geq r$)

$(w - r + i)$ Ones from E' cover $(w - r + i)$ out of w Ones of $(E + E')$

i Ones from E' cover i out of $(m-w)$ Zeroes of $(E + E')$

to give r ones in E.

This can be done in

$$X_{w,r} = \sum_i C^{w}_{w-r+i} \cdot C^{m-w}_{i} \text{ ways:} \qquad \text{(VIII.3)}$$

where i is chosen to satisfy:

$$0 \leq w-r+i \leq w$$

$$0 \leq i \leq m-w$$

$$w-r+2i < \frac{d}{2}$$

i.e.

$$\left.\begin{array}{l} i \geq 0 \\ i \leq r \\ i \leq m-w \\ i < \dfrac{\frac{d}{2} - (w-r)}{2} \end{array}\right\} \qquad \text{(VIII.4)}$$

Case 3 ($w = 0$)

r Ones from E' cover r out of m Zeroes of $(E + E')$

$$X_{w,r} = C^{m}_{r} \qquad \text{(VIII.5)}$$

$$0 \leq r < \frac{d}{2} \qquad \text{(VIII.6)}$$

Case 4 ($w = m$)

$(m - r)$ Ones from E' cover $(m - r)$ out of m Ones of $(E + E')$

$$X_{w,r} = C^{m}_{m-r} \qquad \text{(VIII.7)}$$

$$0 \leq m-r < \frac{d}{2} \qquad \text{(VIII.8)}$$

These relations can be evaluated easily, by digital computer. Where values of w and r do not satisfy relations (VIII.2), (VIII.4), (VIII.6), or (VIII.8), X is zero.

Conversely, $X_{w,r}$ different original error words, of weight r, (out of a total of C_r^m such words) produce a particular corrected error word of weight, w, under the process of Golay decoding by the addition of the coset leader (i.e. the error word of lowest weight, producing the same syndrome as the actual error word).

APPENDIX IX

REPLICATION OF m-BIT DATA WORD

Consider a codeword of m elements. The error probability of each element is P. We take n replicates of the same transmitted codeword. We choose the most numerous result (or choose at random, where two or more results are equally numerous). The probability of getting a codeword with t elements in error is:

$$P_t = C_t^m \cdot P^t \cdot (1-P)^{m-t}$$

The probability of getting r replicates with no errors, and u replicates with t errors, and (n - r - u) replicates with any other number of errors, is:

$$P_{(r,u,t)} = C_r^n . P_o^r . C_u^{n-r} . P_t^u . C_{n-r-u}^{n-r-u} . (1-P_o-P_t)^{n-r-u}$$

The probability that s of these u, have a particular arrangement of the t errors is:

$$P_{(s,u,t)} = C_s^u \cdot \left(\frac{1}{C_t^m}\right)^s \cdot \left(1 - \frac{1}{C_t^m}\right)^{u-s}$$

Thus, the probability of getting r replicates with no errors, and s replicates with one particular arrangement of t errors is:

$$P_{(r,s,t)} = \sum_{u=s}^{n-r} P_{(r,u,t)} \cdot P_{(s,u,t)}$$

$$= \frac{n!}{r!\ s!} . (1-P)^{m.r} . \sum_{u=s}^{n-r} \frac{(C_t^m-1)^{u-s} . P^{t.u} . (1-P)^{u(m-t)} . (1-P_o-P_t)^{n-r-u}}{(n-r-u)!\ (u-s)!} \tag{IX.1}$$

If we sum this over all t, and all arrangements of t element errors, we get the probability P(r,s) that the most numerous replicated error word is of order, s, plus a slight excess probability due to the multiple inclusion (instead of once) of cases where more than one arrangement is of order, s, and due to the inclusion of cases where other arrangements are of order greater than s (which should not be included at all). If $s \geqslant 2$, these errors should be small (or zero, if n is small).

$$P_{(r,s)} \approx \sum_{t=1}^{m} C_t^m \cdot P_{(r,s,t)} \quad ; \quad S \geqslant 2 \tag{IX.2}$$

If $s = 0$, there is no possibility of error. If $s = 1$, and $r > 1$, there is also no probability of error. Therefore the only other cases requiring consideration are $s = 1$, $r = 1$, where the probability of error is $(n - 1)/n$; and $s = 1$, $r = 0$, where the probability is unity.

$$P_{(1,1)} = C_1^n . P_o^1 . (1-P_o)^{n-1} - \sum_{s=2}^{n-1} P_{(1,s)} \tag{IX.3}$$

$$P_{(0,1)} = C_o^n . P_o^o (1-P_o)^n - \sum_{s=2}^{n} P_{(o,s)} \tag{IX.4}$$

Ignoring multiple occurrences of order s, replicates, the probability of error is unity for $s > r$, and ½ for $s = r$, other than $s = r = 1$. By combining this with equations (IX.2) (IX.3) (IX.4) we find the approximate relation for data word error probability as:

$$P_D = C_o^n . P_o^o . (1-P_o)^n + \frac{n-1}{n} . C_1^n . P_o^1 . (1-P_o)^{n-1} + \sum_{r=1}^{[n/2]} \sum_{s=2}^{n-r} P_{(r,s)} . Q_{(r,s)} \tag{IX.5}$$

where
$$Q_{(r,s)} = \begin{cases} 1 & \text{if } 1 < r < s \\ 0 & \text{if } r > s \\ \frac{1}{2} & \text{if } r = s \\ 1/n & \text{if } r = 1 \end{cases}$$

TABLE 1. DECODER PERFORMANCES, ONE OUT OF (K + 1) CODING LINEAR ENVELOPE DETECTION

(K + 1)	Word error probability	Required signal to noise ratio (dB)				
		Maximum amplitude decoder	Element threshold decoder			
			$P_A = P_B$		Optimum threshold	
	(P_w)	(P_{w1})	P_{w2}	P_{w3}	P_{w2}	P_{w3}
2	10^{-2}	9.0	12.1	11.3	12.1	11.3
	10^{-3}	11.0	14.0	12.5	14.0	12.5
	10^{-4}	12.3	15.5	15.1	15.5	15.1
	10^{-5}	13.3				
4	10^{-2}	9.9	12.8	12.2	12.6	12.0
	10^{-3}	11.7	14.5	14.1	14.4	14.0
	10^{-4}	12.8	15.8	15.5	15.7	15.4
	10^{-5}	13.8				
8	10^{-2}	10.5	13.3	12.8	12.8	12.3
	10^{-3}	12.1	14.9	14.5	14.6	14.2
	10^{-4}	13.2	16.1	15.8	15.8	15.5
	10^{-5}	14.1				
16	10^{-2}	11.0	13.8	13.3		
	10^{-3}	12.4	15.3	15.0		
	10^{-4}	13.5	16.4	16.2		
	10^{-5}	14.3				

TABLE 2. POST DECODING INTEGRATION PERFORMANCE. ONE OUT OF (K + 1) CODING. MAXIMUM AMPLITUDE DETECTION. LINEAR ENVELOPE DETECTION

K+1 \ P_{wn} \ n		Gain (dB)						Integration loss (dB)					
		2	3	4	5	7	10	2	3	4	5	7	10
2	10^{-2}	0.0	2.5	2.5	3.9	-	-	3.0	2.3	2.3	3.1	-	-
	10^{-3}	0.0	2.7	2.7	4.2	5.1	5.9	3.0	2.1	2.1	2.8	3.3	4.1
	10^{-4}	0.0	2.7	2.7	4.3	5.4	6.1	3.0	2.0	2.0	2.7	3.1	3.9
4	10^{-2}	0.0	2.4	2.9	3.8	4.8	-	3.0	2.4	3.1	3.2	3.7	-
	10^{-3}	0.0	2.5	2.9	3.9	5.1	6.2	3.0	2.3	3.1	3.1	3.4	3.8
	10^{-4}	0.0	2.5	2.9	4.0	5.1	6.3	3.0	2.2	3.1	3.0	3.4	3.7
8	10^{-2}	0.0	2.1	3.0	3.8	4.5	5.2	3.0	2.7	3.0	3.2	3.9	4.8
	10^{-3}	0.0	2.3	3.1	3.8	4.9	5.8	3.0	2.4	2.9	3.2	3.6	4.2
	10^{-4}	0.0	2.4	3.0	4.0	5.0	6.0	3.0	2.4	3.0	3.0	3.5	4.0
16	10^{-2}	0.0	2.0	3.0	3.7	4.1	4.7	3.0	2.8	3.0	3.3	4.3	5.3
	10^{-3}	0.0	2.2	3.1	3.8	4.5	5.2	3.0	2.6	2.9	3.2	4.0	4.8
	10^{-4}	0.0	2.2	3.2	4.0	4.8	5.6	3.0	2.5	2.8	3.0	3.7	4.4

TABLE 3. POST DETECTION INTEGRATION PERFORMANCE. ONE OUT OF (K + 1) CODING. THRESHOLD DETECTION. LINEAR ENVELOPE DETECTION

K+1 \ P_{wn} \ n		Gain (dB)						Integration loss (dB)					
		2	3	4	5	7	10	2	3	4	5	7	10
2	10^{-2}	2.8	4.4	5.4	6.3	7.4	8.1	0.2	0.4	0.6	0.7	1.0	1.9
	10^{-3}	2.9	4.6	5.7	6.5	7.7	8.8	0.1	0.2	0.3	0.5	0.7	1.2
	10^{-4}	3.2	4.7	5.9	6.8	8.0	9.3	-0.2	0.0	0.1	0.2	0.5	0.7
4	10^{-2}	2.7	4.3	5.4	6.2	7.2	8.7	0.3	0.5	0.6	0.8	1.2	1.3
	10^{-3}	2.9	4.4	5.7	6.4	7.6	8.9	0.1	0.3	0.3	0.5	0.8	1.1
	10^{-4}	3.1	4.7	5.8	6.7	8.0	9.2	-0.1	0.1	0.2	0.3	0.4	0.7
8	10^{-2}	2.6	4.3	5.3	6.1	7.3	8.6	0.4	0.5	0.7	0.9	1.1	1.4
	10^{-3}	2.8	4.4	5.6	6.5	7.7	8.7	0.2	0.4	0.4	0.5	0.7	1.2
	10^{-4}	3.0	4.7	5.8	6.7	8.0	9.2	0.0	0.1	0.2	0.2	0.4	0.8
16	10^{-2}	2.5	4.3	5.4	6.2	7.4	8.5	0.5	0.5	0.6	0.8	1.1	1.5
	10^{-3}	2.7	4.4	5.6	6.5	7.6	9.0	0.3	0.4	0.4	0.5	0.8	1.0
	10^{-4}	3.0	4.7	5.8	6.8	8.1	9.4	0.0	0.1	0.2	0.2	0.3	0.6

TABLE 4. Xw,r AS A FUNCTION OF w AND r

r \ w	0	7	8	11	12	15	16	23
0	1	0	0	0	0	0	0	0
1	23	0	0	0	0	0	0	0
2	253	0	0	0	0	0	0	0
3	1771	0	0	0	0	0	0	0
4	0	35	0	0	0	0	0	0
5	0	21	56	0	0	0	0	0
6	0	343	28	0	0	0	0	0
7	0	113	428	0	0	0	0	0
8	0	856	121	165	0	0	0	0
9	0	120	855	55	220	0	0	0
10	0	560	105	671	66	0	0	0
11	0	0	455	133	738	0	0	0
12	0	0	0	738	133	455	0	0
13	0	0	0	66	671	105	560	0
14	0	0	0	220	55	855	120	0
15	0	0	0	0	165	121	856	0
16	0	0	0	0	0	428	113	0
17	0	0	0	0	0	28	343	0
18	0	0	0	0	0	56	21	0
19	0	0	0	0	0	0	35	0
20	0	0	0	0	0	0	0	1771
21	0	0	0	0	0	0	0	253
22	0	0	0	0	0	0	0	23
23	0	0	0	0	0	0	0	1
$\sum_r$	2048	2048	2048	2048	2048	2048	2048	2048

TABLE 5. LOG_{10} Pw AS A FUNCTION OF w, AND LOG_{10} Pe

logPe \ w	0	7	8	11	12	15	16	23
-0.50	-1.42	-0.49	-0.43	-0.78	-1.04	-2.26	-2.83	-7.22
-0.75	-0.40	-0.44	-0.65	-1.84	-2.40	-4.52	-5.40	-11.99
-1.00	-0.09	-0.84	-1.31	-3.32	-4.16	-7.12	-8.29	-16.88
-1.25	-0.02	-1.50	-2.23	-5.03	-6.14	-9.91	-11.34	-21.82
-1.50	-0.00	-2.31	-3.29	-6.88	-8.24	-12.79	-14.48	-26.79
-2.00	-0.00	-4.13	-5.62	-10.74	-12.61	-18.68	-20.89	-36.76
-2.50	-0.00	-6.08	-8.07	-14.69	-17.07	-24.65	-27.36	-46.76
-3.00	-0.00	-8.06	-10.56	-18.68	-21.55	-30.64	-33.85	-56.75
Nw	1	253	506	1288	1288	506	253	1

TABLE 6. REPLICATION ERROR WITH POST GOLAY DECODING INTEGRATION

n	x_o	Distribution between remaining x	Probability of getting assumed distribution	Probability of getting a decoding error
2	1	1	$2.P_o.\sum_{7}^{23} Pw = 2P_o.z$	$\frac{1}{2}$
	0	Any	z^2	1
3	1	2	$3\ P_o.\Sigma\ Pw^2$	$^2/_3 + {}^1/_3\ Nw$
	1	1,1	$6\ P_o\ \Sigma.Pw.Pw$	$^2/_3$
	0	Any	z^3	1
4	2	2	$6\ P_o^2.\Sigma Pw^2$	$^1/_2\ Nw$
	2	1,1		0
	1	3	$4\ P_o\Sigma Pw^3$	$^3/_4 + \frac{3Nw-2}{4Nw^2}$
	1	2,1	$12\ P_o\Sigma Pw^2.Pw$	$^3/_4 + {}^1/_4\ Nw$
	1	1,1,1	$24\ P_o\Sigma Pw.Pw.Pw$	$^3/_4$
	0	Any	z^4	1
5	2	3	$10\ P_o^2\Sigma Pw^3$	$^3/_{2Nw} - {}^1/_{2Nw^2}$
	2	2,1	$30\ P_o^2\Sigma Pw^2.Pw$	$^1/_{2Nw}$
	2	1,1,1		0
	1	4	$5\ P_o\Sigma Pw^4$	$^4/_5 + {}^1/_5\left\{1-\frac{(N-1)(N-2)(N-3)}{N^3}\right\}$
	1	3,1	$20\ P_o\Sigma Pw^3.Pw$	$^4/_5 + {}^1/_5\left\{\frac{3N-2}{N^2}\right\}$
	1	2,2	$30\ P_o\Sigma Pw^2.Pw^2$	$^4/_5 + {}^1/_5\left\{\frac{1}{N_1}+\frac{1}{N_2}-\frac{1}{N_1.N_2}\right\}$
	1	2,1,1	$60\ P_o\Sigma Pw^2.Pw.Pw$	$^4/_5 + {}^1/_5.\frac{1}{Nw}$
	1	1,1,1,1	$120\ P_o\Sigma Pw.Pw.Pw.Pw$	$^4/_5$
	0	Any	z^5	1

TABLE 7. POST GOLAY DECODING INTEGRATION PERFORMANCE

P_{G2} \ n	Gain (dB)				Integration loss (dB)			
	2	3	4	5	2	3	4	5
10^{-2}	0.0	1.3	2.1	2.6	3.0	3.4	3.9	4.4
10^{-3}	0.0	1.6	2.5	3.0	3.0	3.2	3.5	4.0
10^{-4}	0.0	1.7	2.6	3.2	3.0	3.1	3.4	3.8

TABLE 8. DISTRIBUTION OF CORRECTION WEIGHT (E′) AS A FUNCTION OF BIT ERRORS (r) IN GOLAY WORD

r	w	i	E′	Number of cases	Probability (given r)
0	0	0	0	1	1
1	0	0	1	23	1
2	0	0	2	253	1
3	0	0	3	1771	1
4	7	0	3	35	1
5	7	0	2	21	21/77
5	8	0	3	56	56/77
6	7	0	1	7	7/371
6	7	1	3	336	336/371
6	8	0	2	28	28/371
7	7	0	0	1	1/541
7	7	1	2	112	112/541
7	8	0	1	8	8/541
7	8	1	3	420	420/541
8	7	0	1	16	16/1142
8	7	1	3	840	840/1142
8	8	0	0	1	1/1142
8	8	1	2	120	120/1142
8	11	0	3	165	165/1142
9	7	0	2	120	120/1250
9	8	0	1	15	15/1250
9	8	1	3	840	840/1250
9	11	0	2	55	55/1250
9	12	0	3	220	220/1250

TABLE 9. MARGINAL PROBABILITY OF GOLAY ERROR CASCADED CODES. NO INTERLEAVING

Bits per return	1	2	3	4
Returns per Golay word	23	12	8	6
Return errors				
1	0	0	0	$^{1}/_{18}$
2	0	$^{5}/_{54}$	$^{39}/_{98}$	$^{448}/_{675}$
3	0	$^{17}/_{27}$	$^{303}/_{343}$	$^{651}/_{675}$
4	1	$^{235}/_{243}$	$^{4775}/_{4802}$	$\frac{151747}{151875}$
$\geqslant 5$	1	1	1	1

TABLE 10. VALUES OF Q(s,t) FOR USE IN EQUATION 40

s	t	Q(s,t)
$\leqslant 3$	$\leqslant 3$	0
0	4,5,6	0
1	4,5	0
1	6	$^{3.5}/_{371} = {}^{1}/_{106}$
2	4	0
2	5	$^{10.5}/_{77} = {}^{3}/_{22}$
2	6	$^{21}/_{371} = {}^{3}/_{53}$
3	4	$\frac{1}{2}$
3	5	$^{49}/_{77} = {}^{7}/_{11}$
3	6	$^{203}/_{371} = {}^{29}/_{53}$
$\leqslant 3$	$\geqslant 7$	0
> 3	> 3	1

TABLE 11. TEXT AND FIGURE REFERENCES TO BOXES OF FIGURE 22

Box (figure 22)	n	(K+1) code	Golay code	Subsection of report	Relevant appendix	Relevant figures	Relevant tables	Remarks
1	1			2.5		2		
1	> 1			4.3		5,6		
2	> 1			4.4		7,8		
6	1	Y		5.2	4	9	1	
6	> 1			6.1		9,13		
8	1	Y		5.4		12	1	
8	> 1	Y		6.3	7	15	3	
9	> 1	Y		6.2	6	14	2	
10	> 1	Y		6.4		-		Inferior to box 9.
12	> 1			9.4		-		Inferior to other boxes.
16	1	N	Y	7.2		16		Random bit errors at box 11.
16	> 1	N	Y	7.4		16	7	Random. No correction weight
16	> 1	N	Y	7.5		16		Random using correction weight
16	1	Y	Y	8.1		17		No interleaving at box 11
16	1	Y	Y	8.2		18		Full interleaving at box 11
16	2	Y	Y	8.4		19		Within group replication
16	> 1	Y	Y	8.3		16		Full interleaving modified by equation (38)
4	≥ 1	N	N	9.2	9	20		Random bit errors
4	1	Y	N	9.3		21		
4	> 1	Y	N	9.3		-		

TABLE 12. CODING SCHEME PERFORMANCE. $P_D = 10^{-4}$. NO REPLICATION

Coding scheme	Golay code	(K+1)	Number of (K+1 words)	Number of elements	Threshold or maximum amplitude detection	Code rate (bits/element)	Duty cycle (activation of elements)	Relative number of active elements	S/N in active element (dB)	Relative transmitted energy (dB)	Figures used in deriving S/N	Remarks
1	N	-	-	12	T	1	0.5	1	16.1	0.0	20,2	No redundancy
2	N	2	12	24	T	½	0.5	2	16.1	+3.0	21,12,2	
2	N	4	6	24	T	½	0.25	1	16.2	+0.1	21,12,2	
2	N	8	4	32	T	3/8	0.125	2/3	16.3	-1.5	21,12,2	
2	N	16	3	48	T	¼	0.0625	½	16.4	-2.7	21,12,2	
3	N	2	12	24	MA	½	0.5	2	13.4	+0.3	21,9	
3	N	4	6	24	MA	½	0.25	1	13.6	-2.5	21,9	
3	N	8	4	32	MA	3/8	0.125	2/3	13.7	-4.1	21,9	
3	N	16	3	48	MA	¼	0.0625	½	13.9	-5.2	21,9	
4	Y	-	-	23	T	12/23	0.5	23/12	11.2	-2.1	16,2	
5	Y	2	24	48	T	¼	0.5	4	11.2	+1.1	17,12,2	
5	Y	4	12	48	T	¼	0.25	2	11.7	-1.4	18,12,2	
5	Y	8	8	64	T	3/16	0.125	4/3	12.2	-2.6	18,12,2	
5	Y	16	6	96	T	1/8	0.0625	1	12.9	-3.2	18,12,2	
6	Y	2	24	48	MA	¼	0.5	4	8.9	-1.2	17,9	Best S/N
6	Y	4	12	48	MA	¼	0.25	2	9.4	-3.7	18,9	Best choice
6	Y	8	8	64	MA	3/16	0.125	4/3	9.9	-4.9	18,9	
6	Y	16	6	96	MA	1/8	0.0625	1	10.4	-5.7	18,9	Best energy

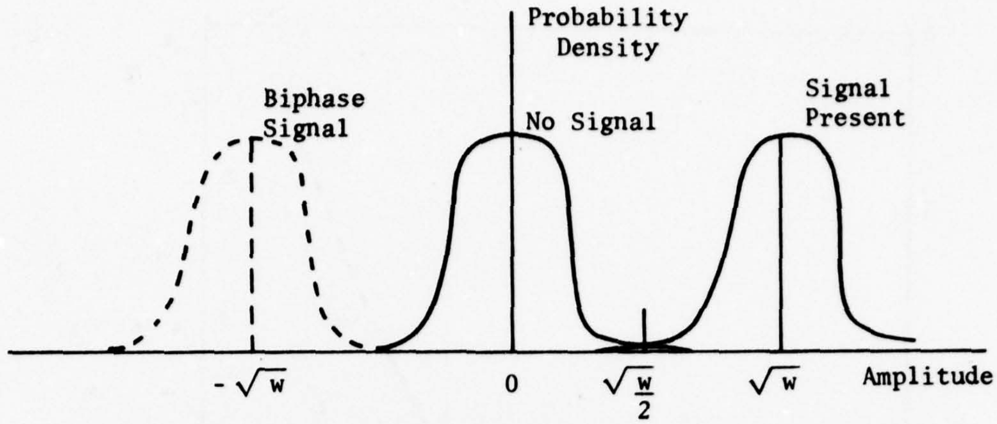


Figure 1. Derived Signal Distributions - Coherent Detection

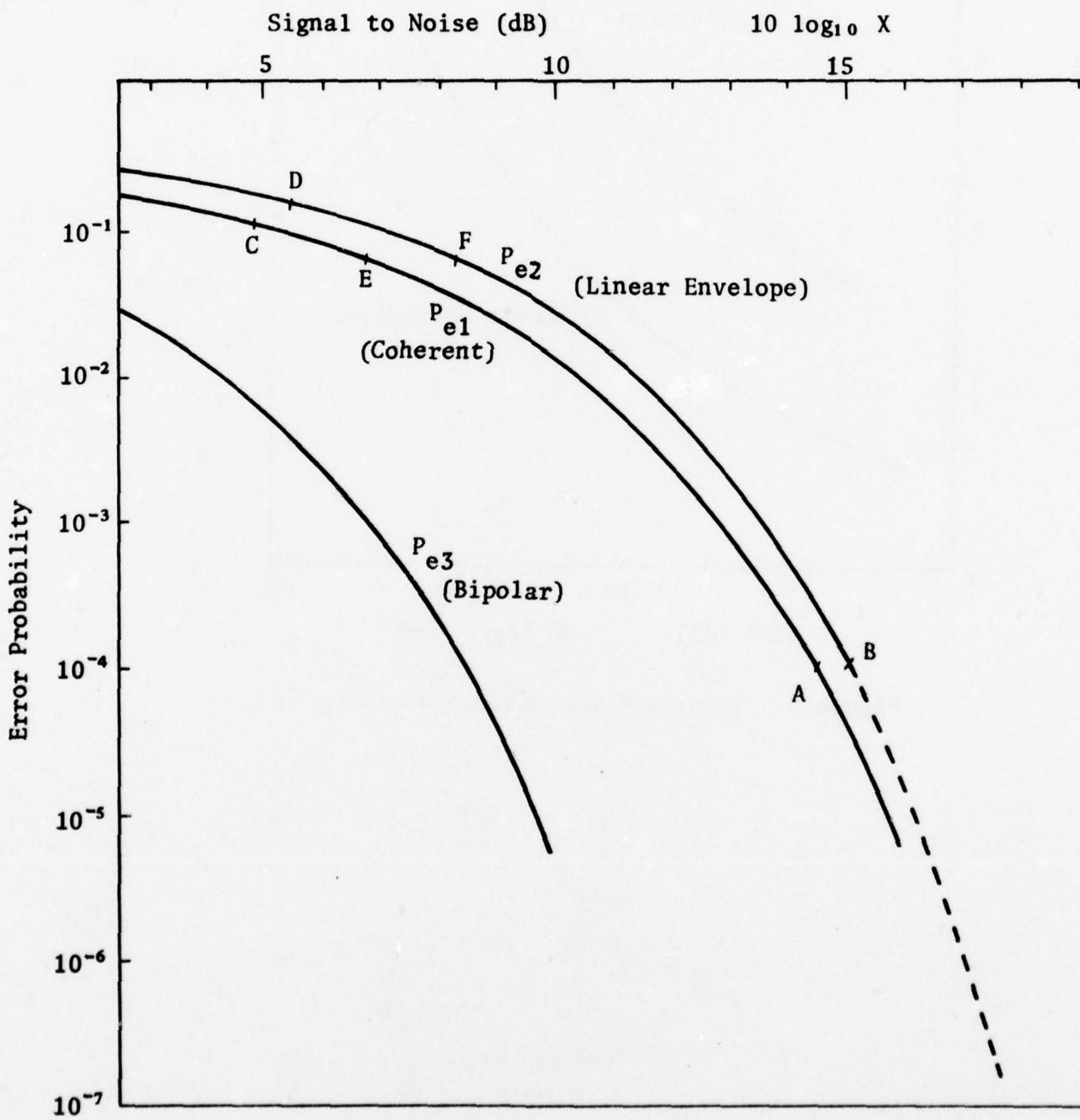


Figure 2. Element Error Probability Versus S/N at Filter Output

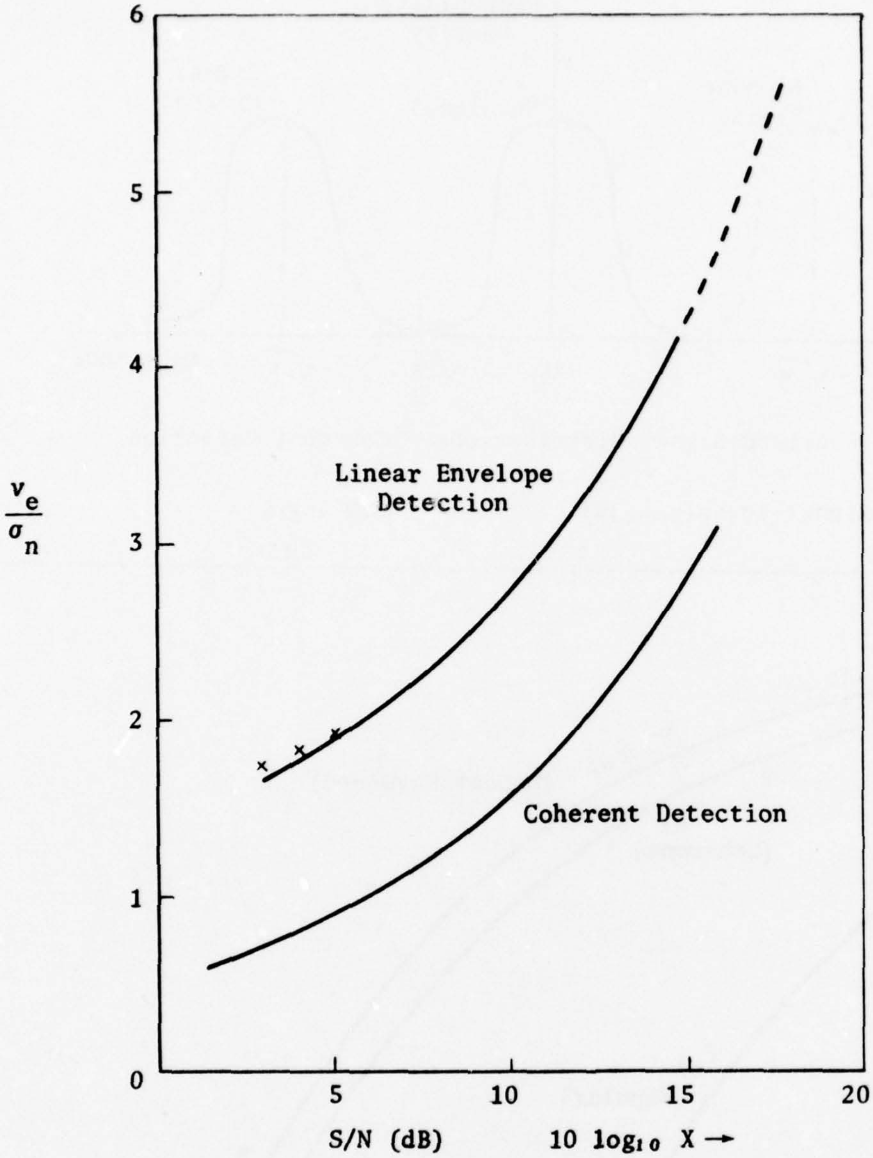


Figure 3. Threshold for Optimum Decoding

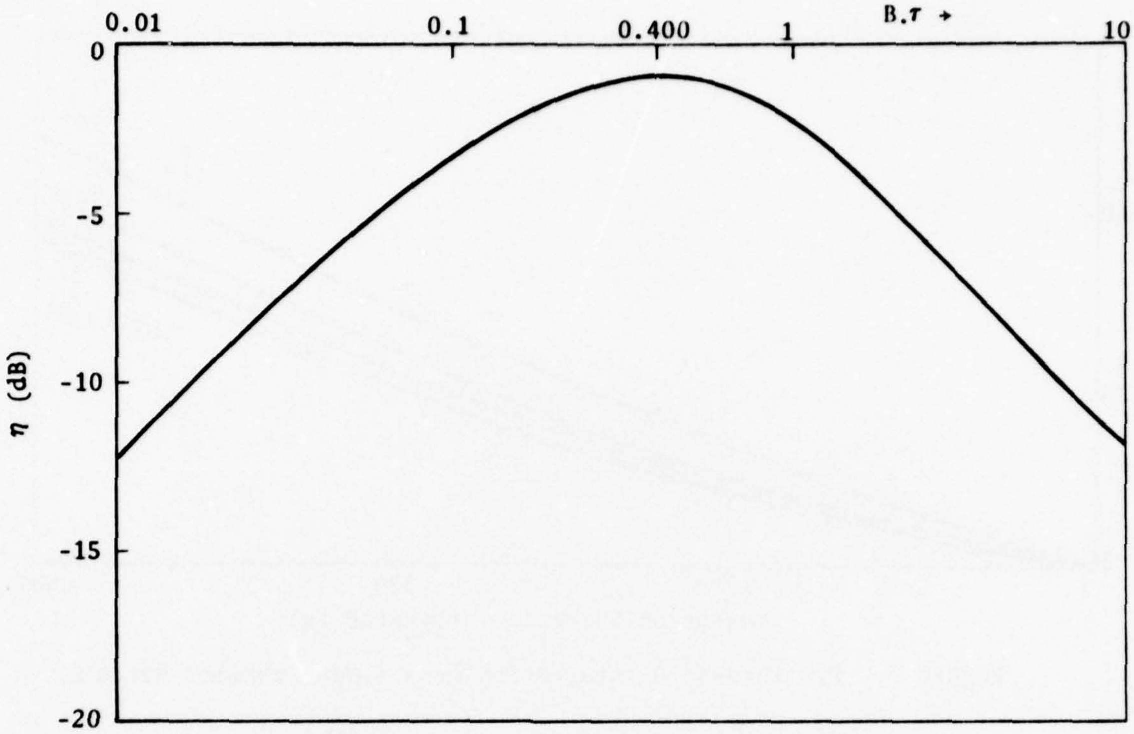


Figure 4. Efficiency of Single Tuned Filter

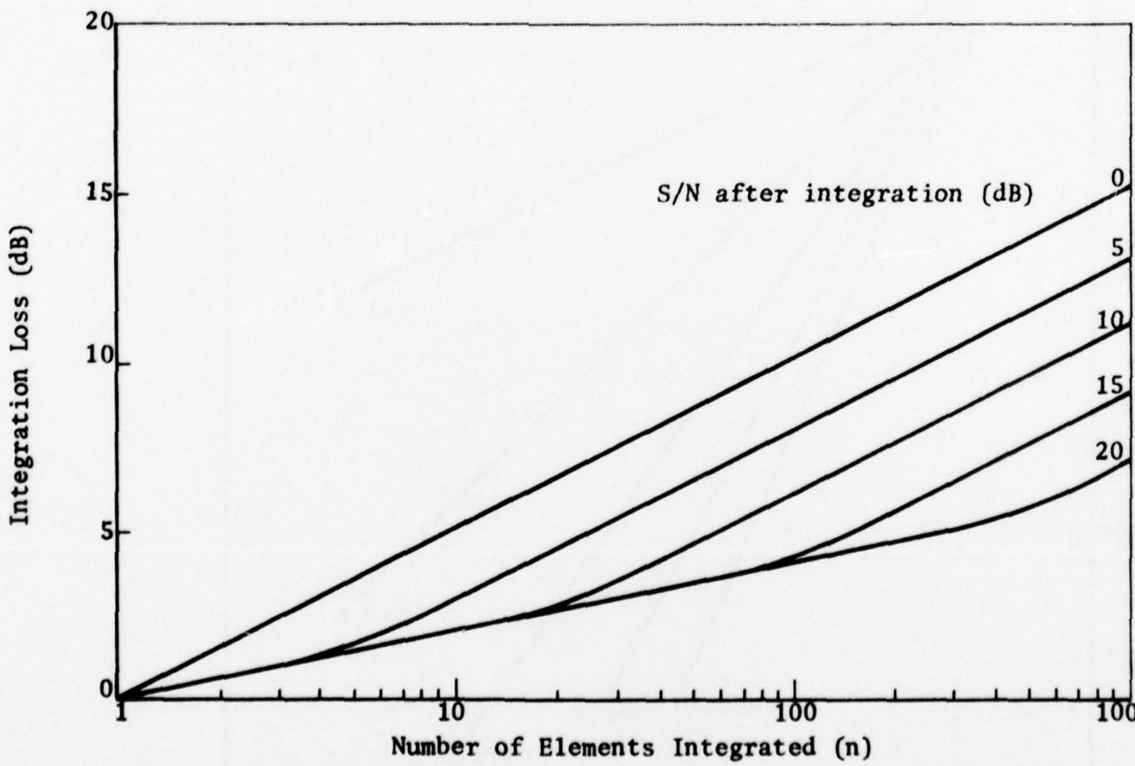


Figure 5. Integration Loss - Non-Coherent Detector

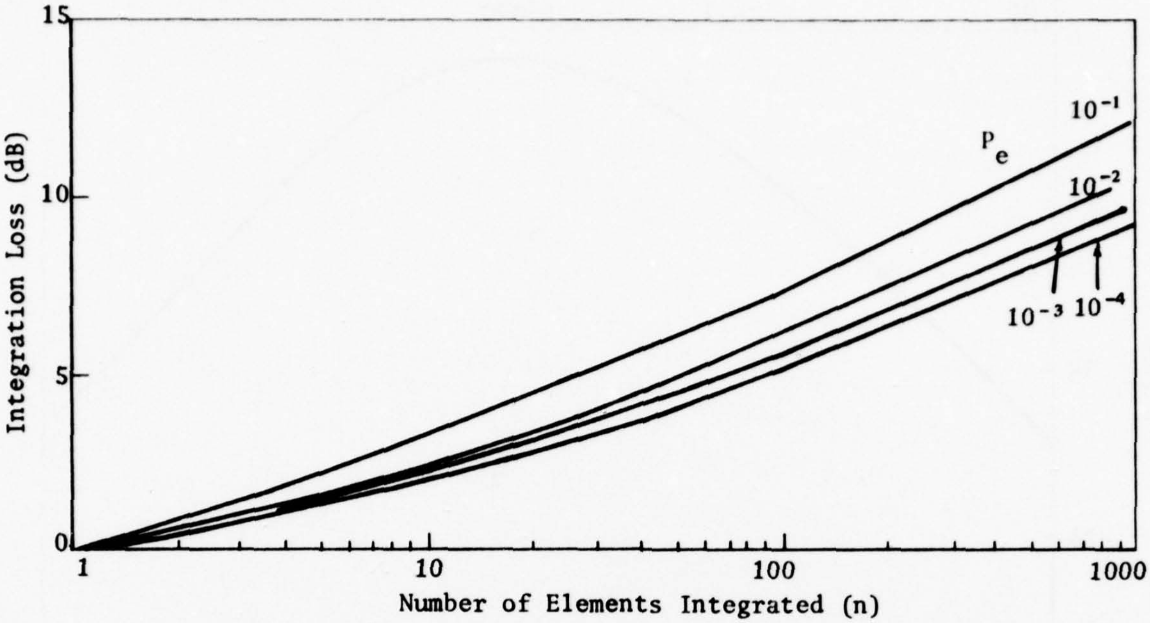


Figure 6. Pre-Threshold Integration Loss - Non-Coherent Detector

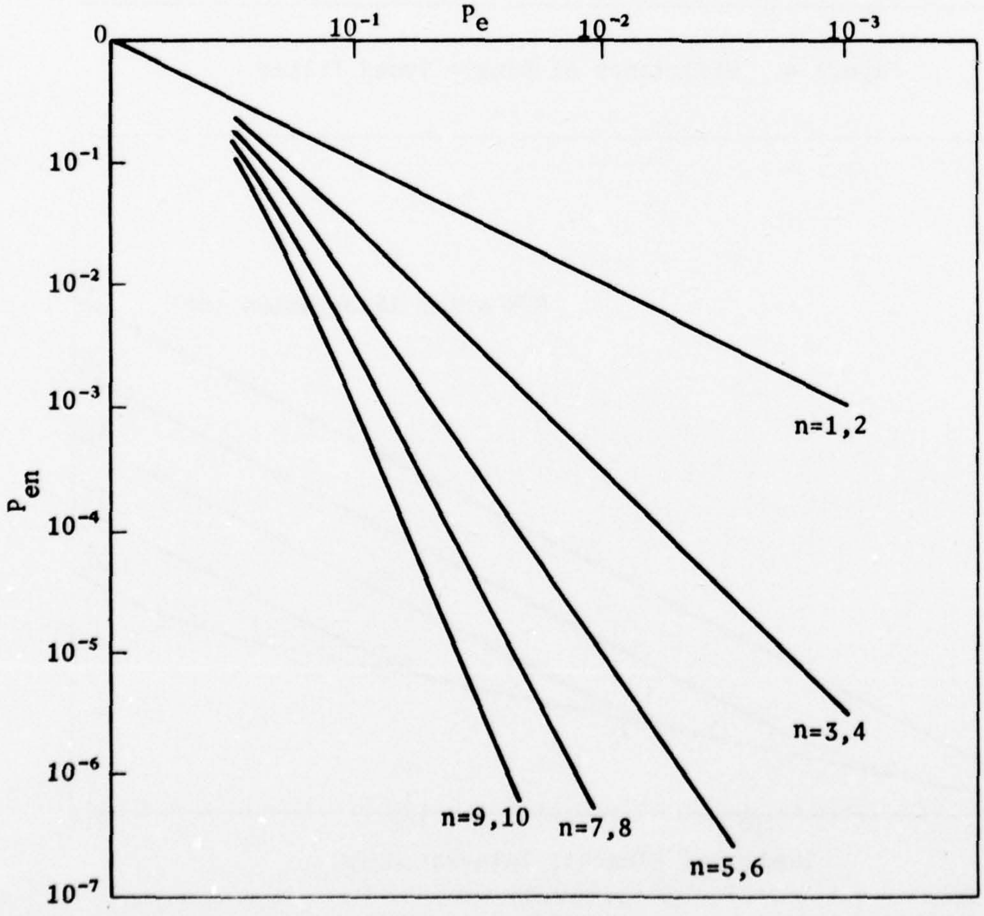


Figure 7. Post Threshold Integration - Single Element

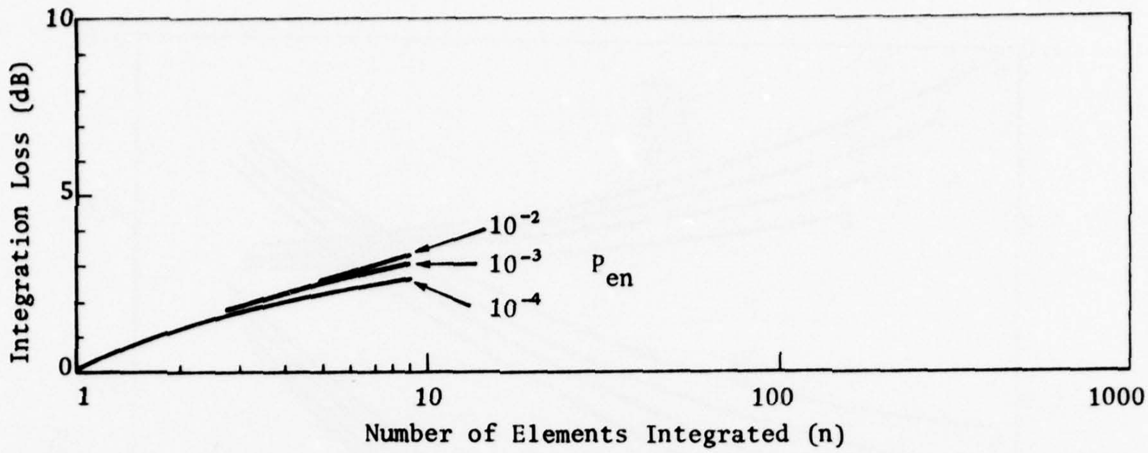


Figure 8. Post Threshold Integration Loss - Single Element

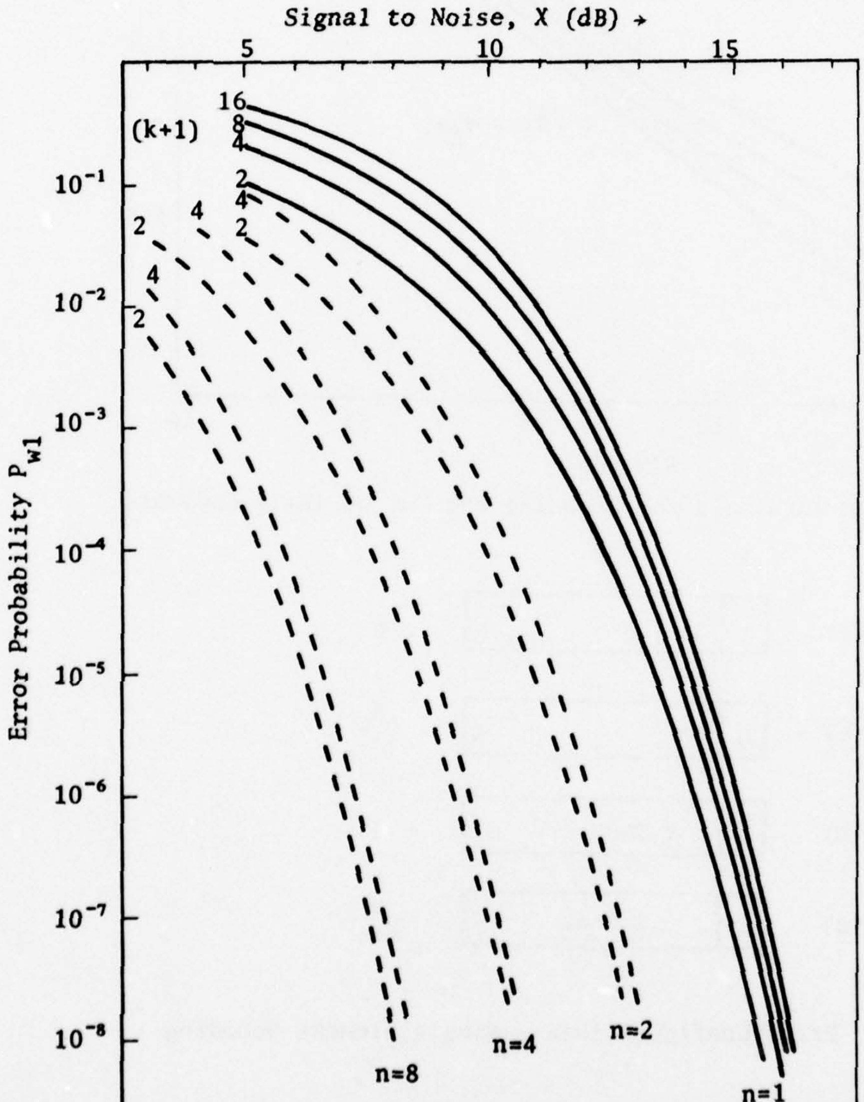


Figure 9. Word Error Probability. One out of (K + 1) Code

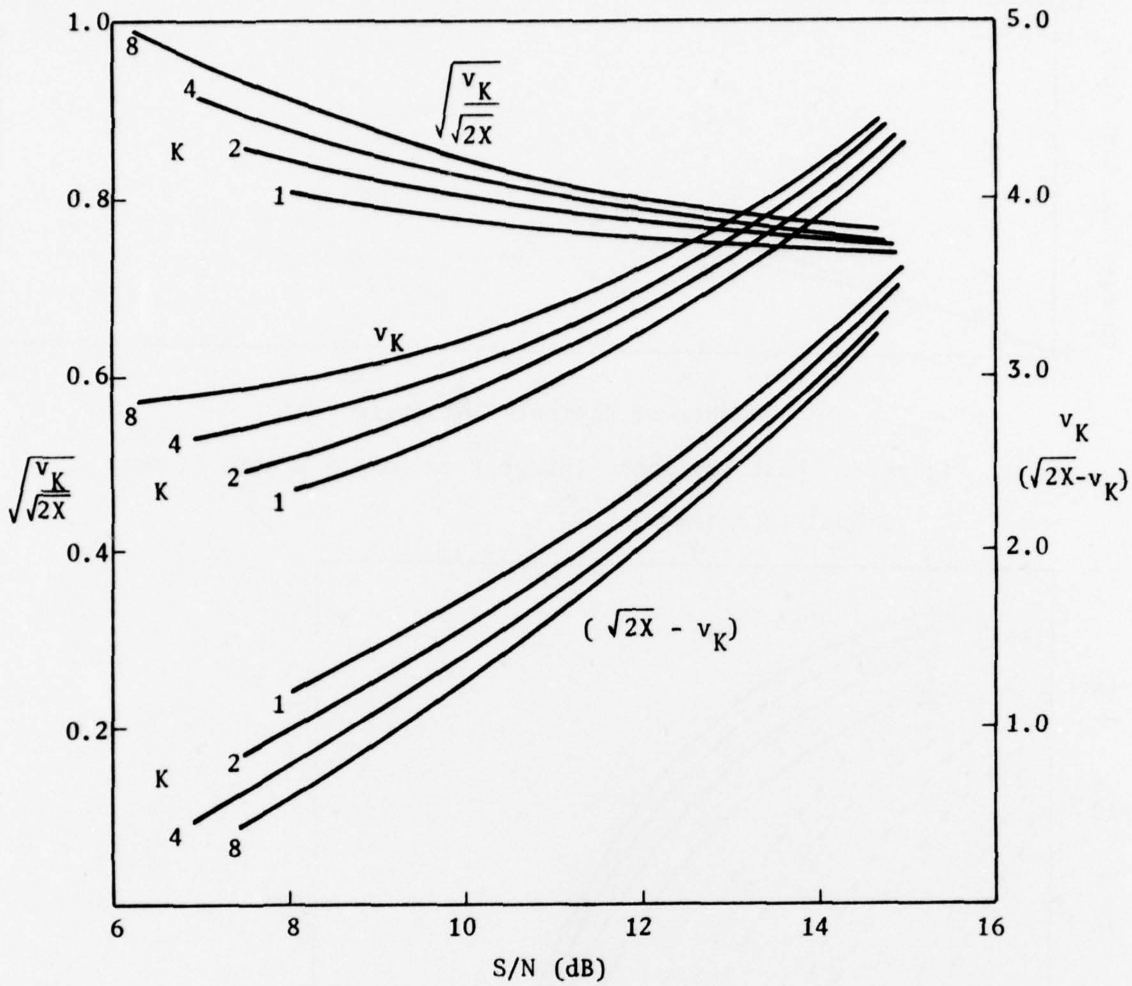


Figure 10. Optimum Threshold for Decoding One Out of (K+1) Codeward

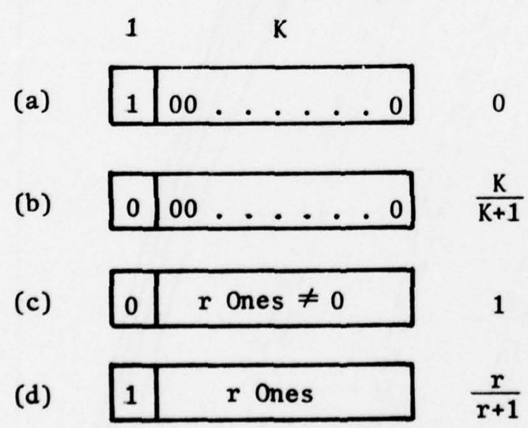


Figure 11. Error Configurations - Single Element Decoding

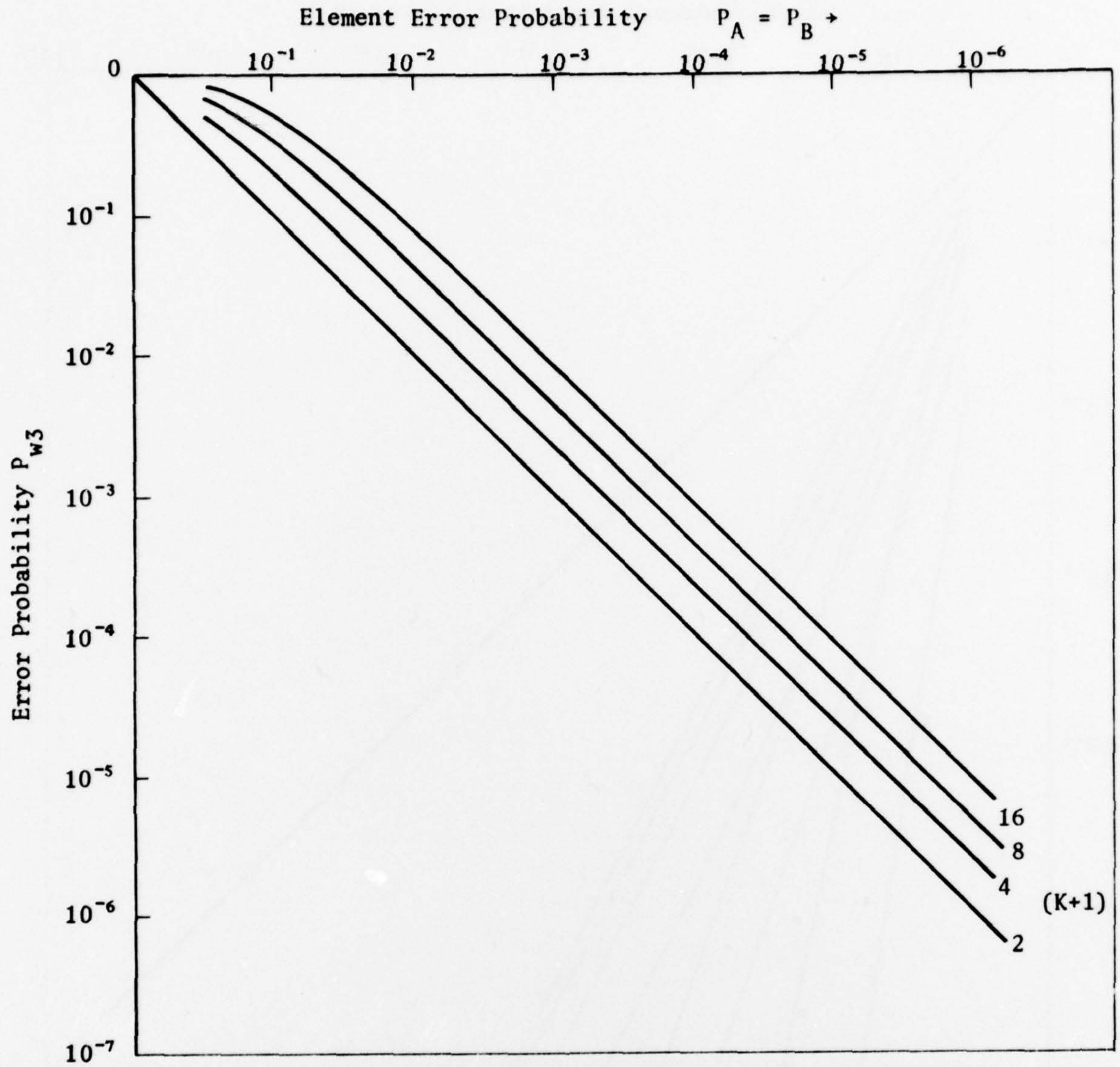


Figure 12. Word Error Probability. One Out of (K+1) Code. Element Threshold ($P_A = P_B$)

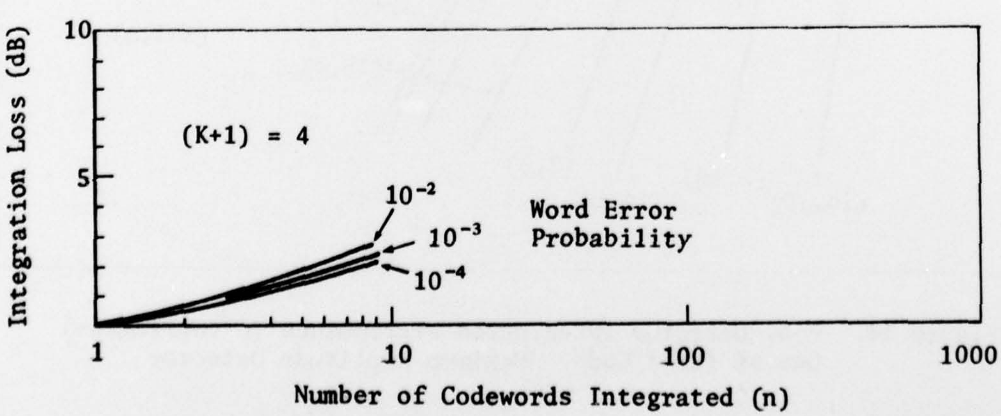


Figure 13. Integration Loss. Maximum Amplitude Decoding

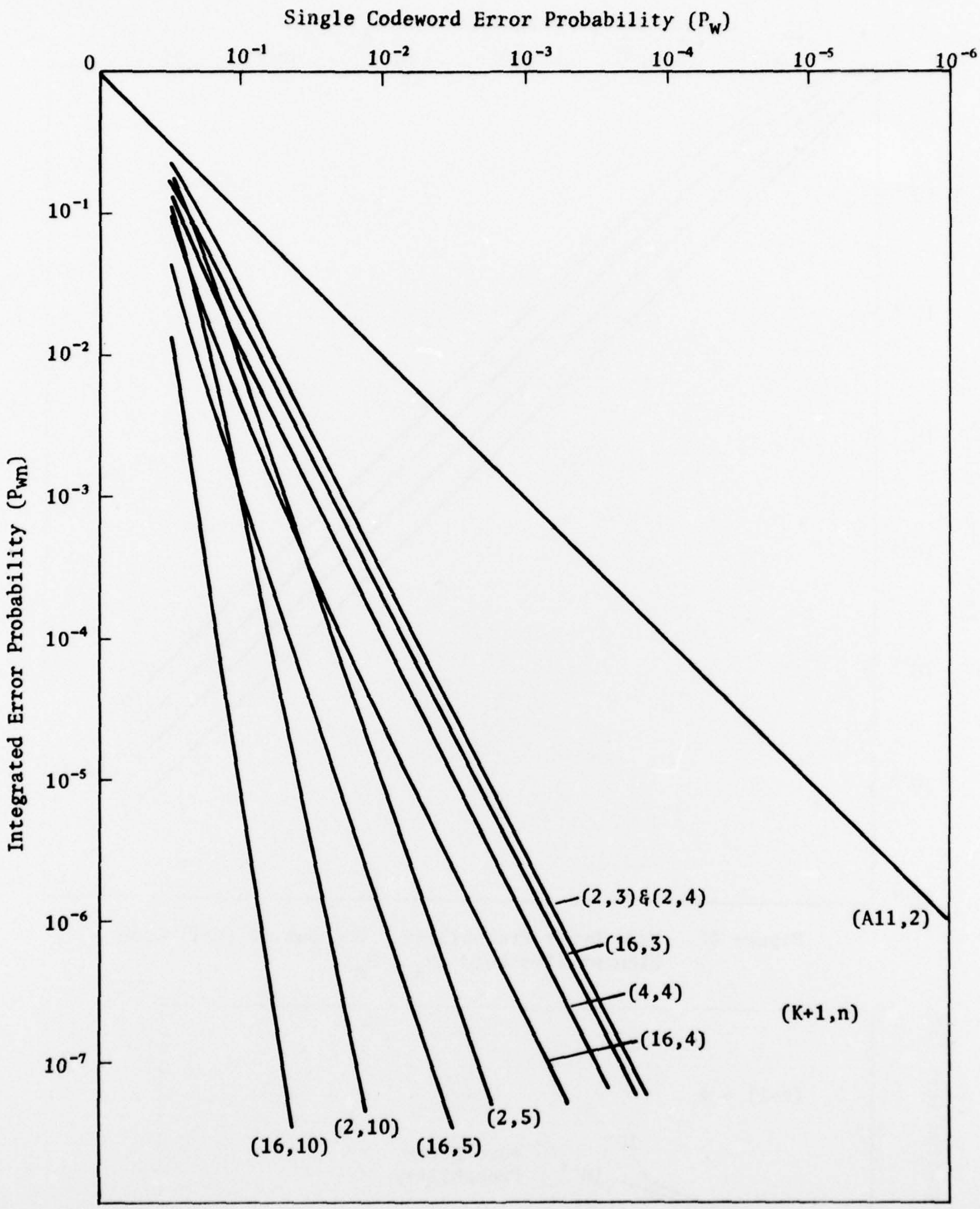


Figure 14. Post Detector Integration Performance (n replicates) One of (K+1) Code. Maximum Amplitude Detector

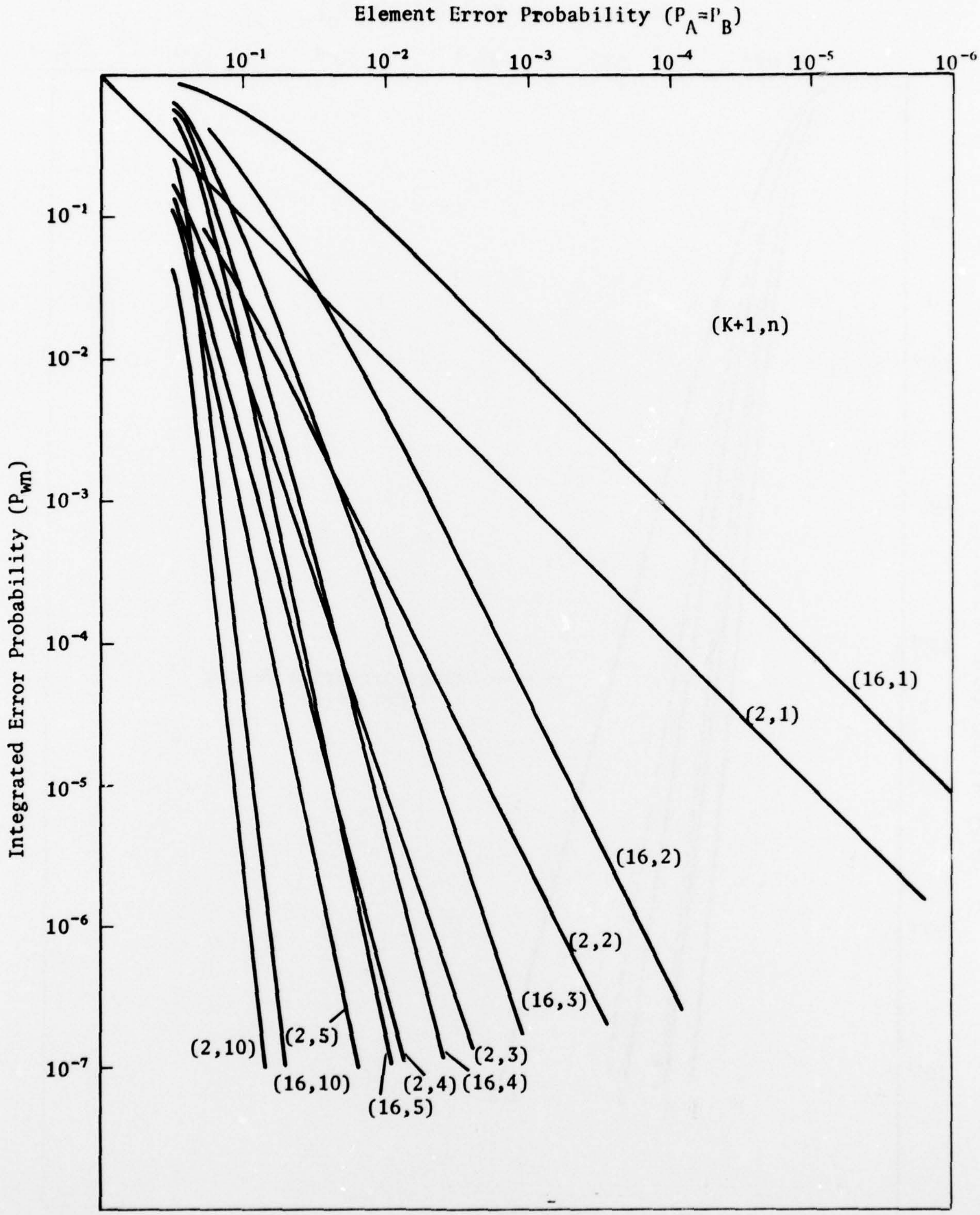


Figure 15. Post Detector Integration Performance (n replicates) (K+1) Code. Summation of Element Threshold Crossings

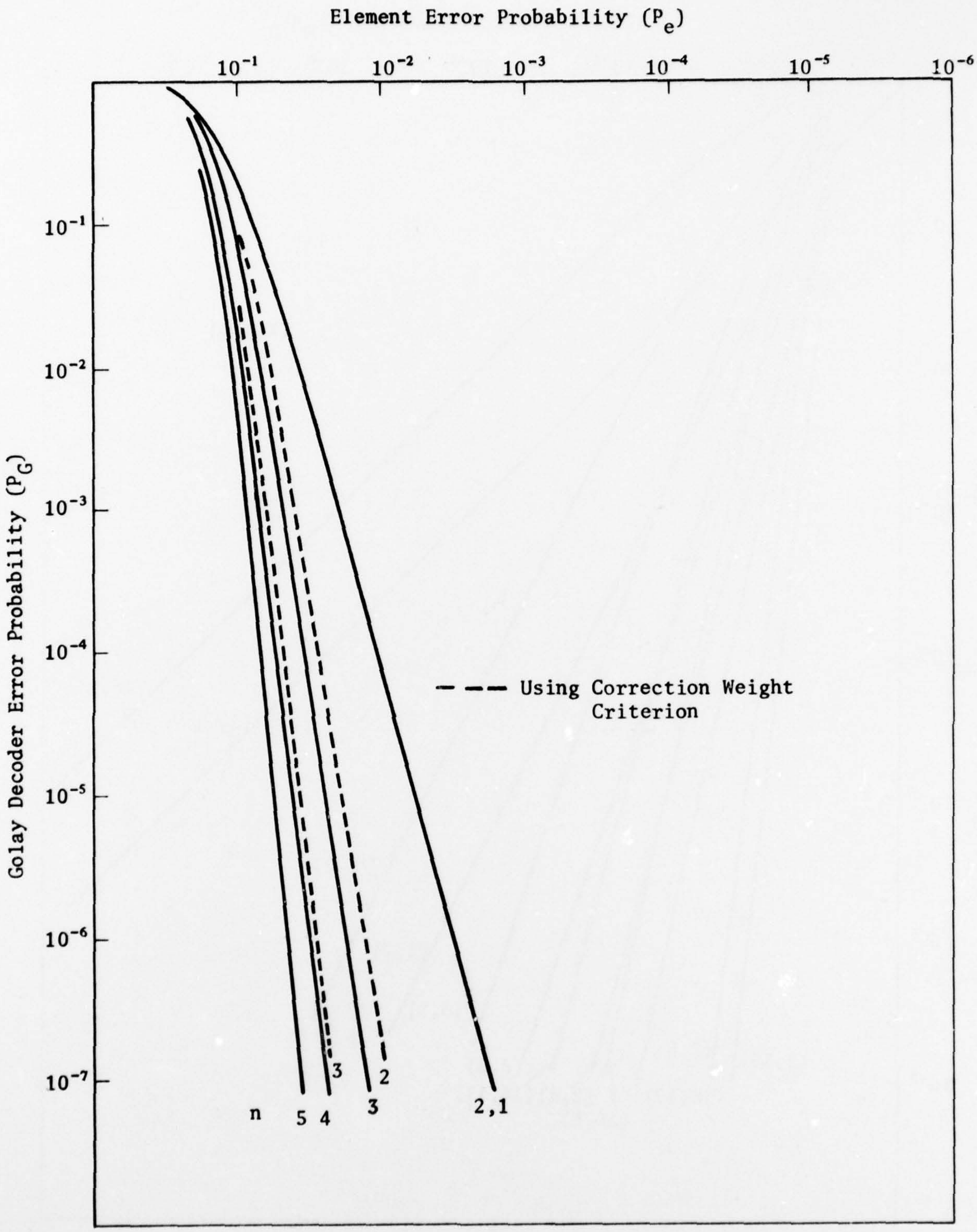


Figure 16. Golay Code Error Probability with Post Decoder Integration

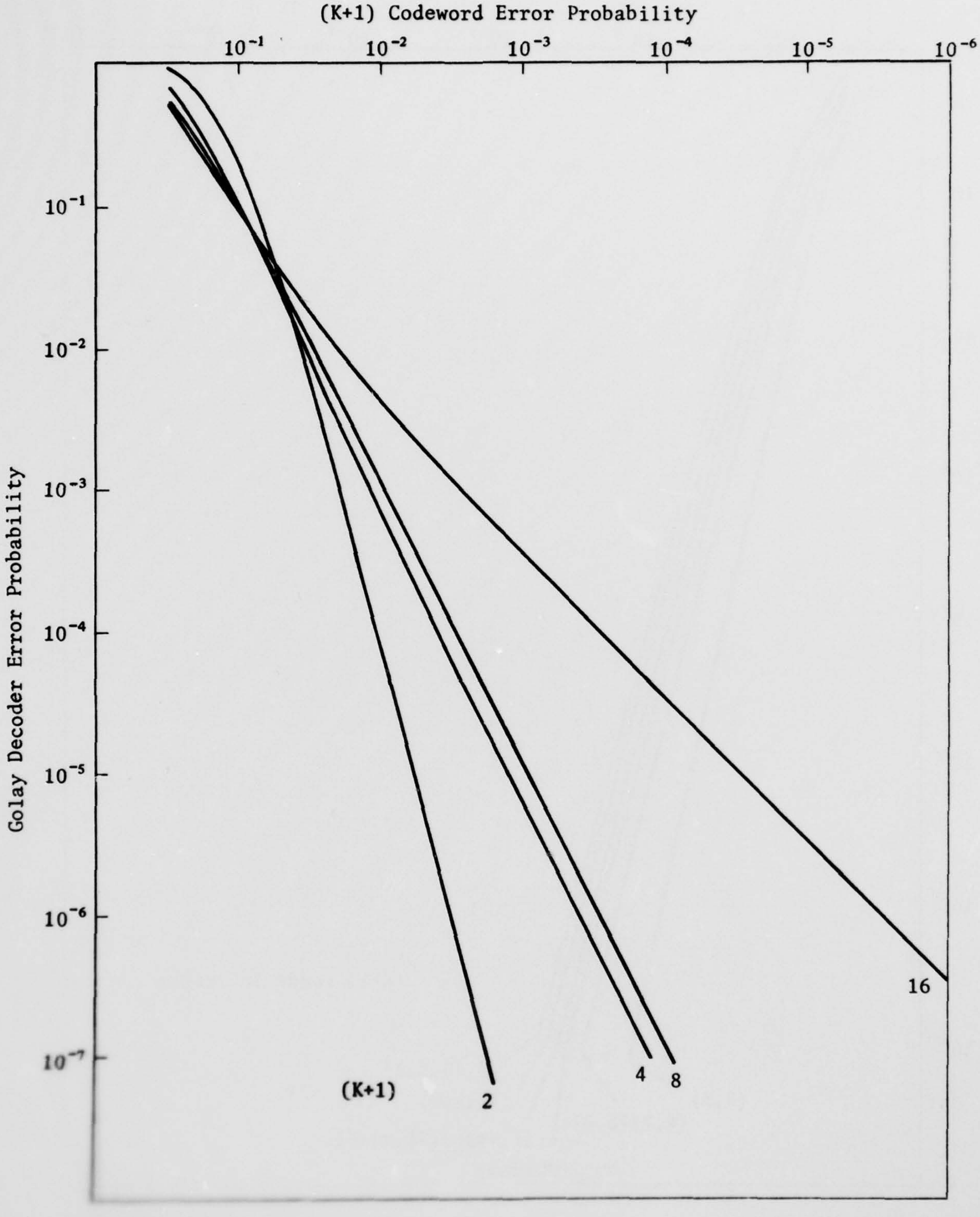


Figure 17. Cascaded (K+1) and Golay Performance. No Interleaving of Bits

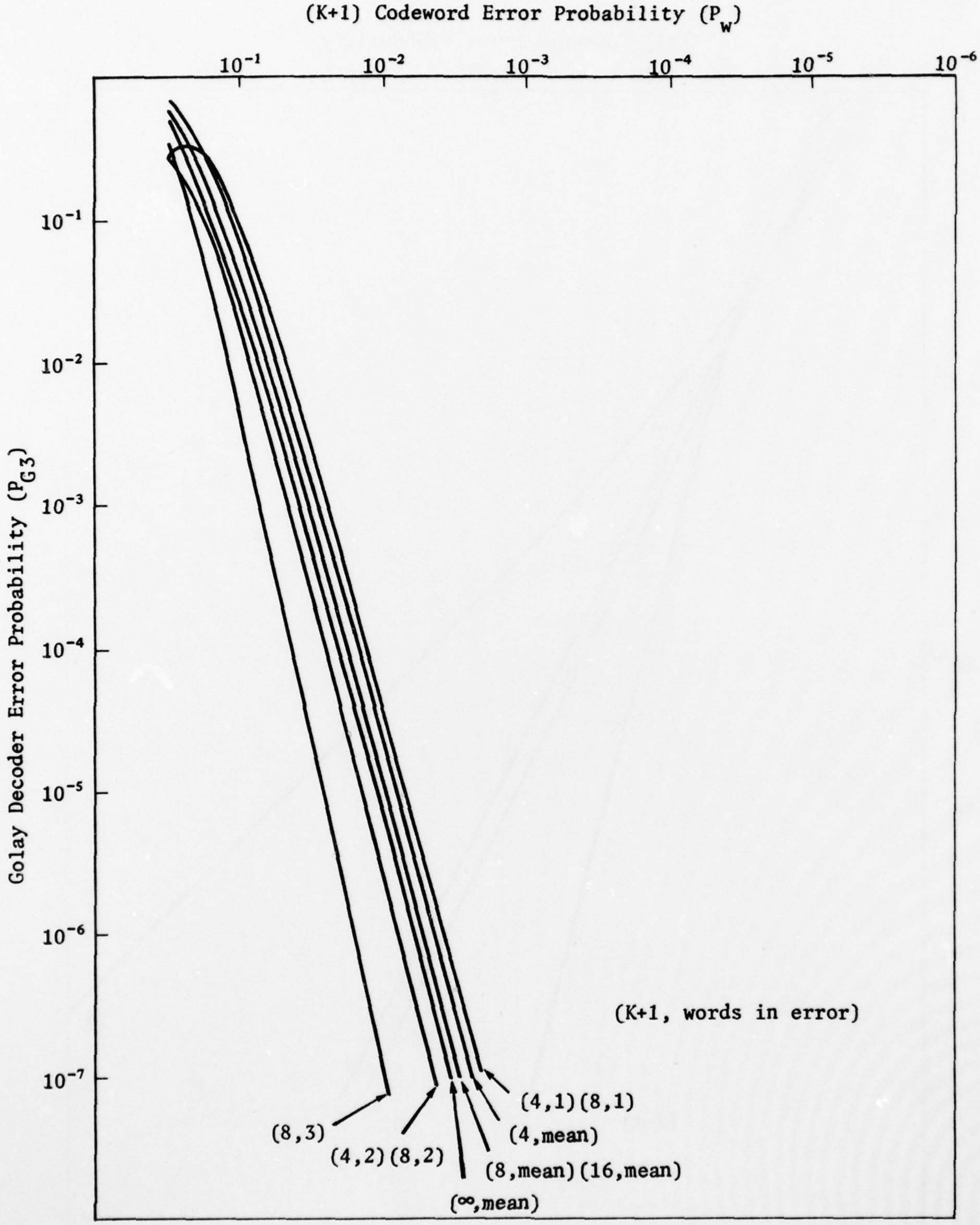


Figure 18. Cascaded (K+1) and Golay Performance. Full Interleaving.

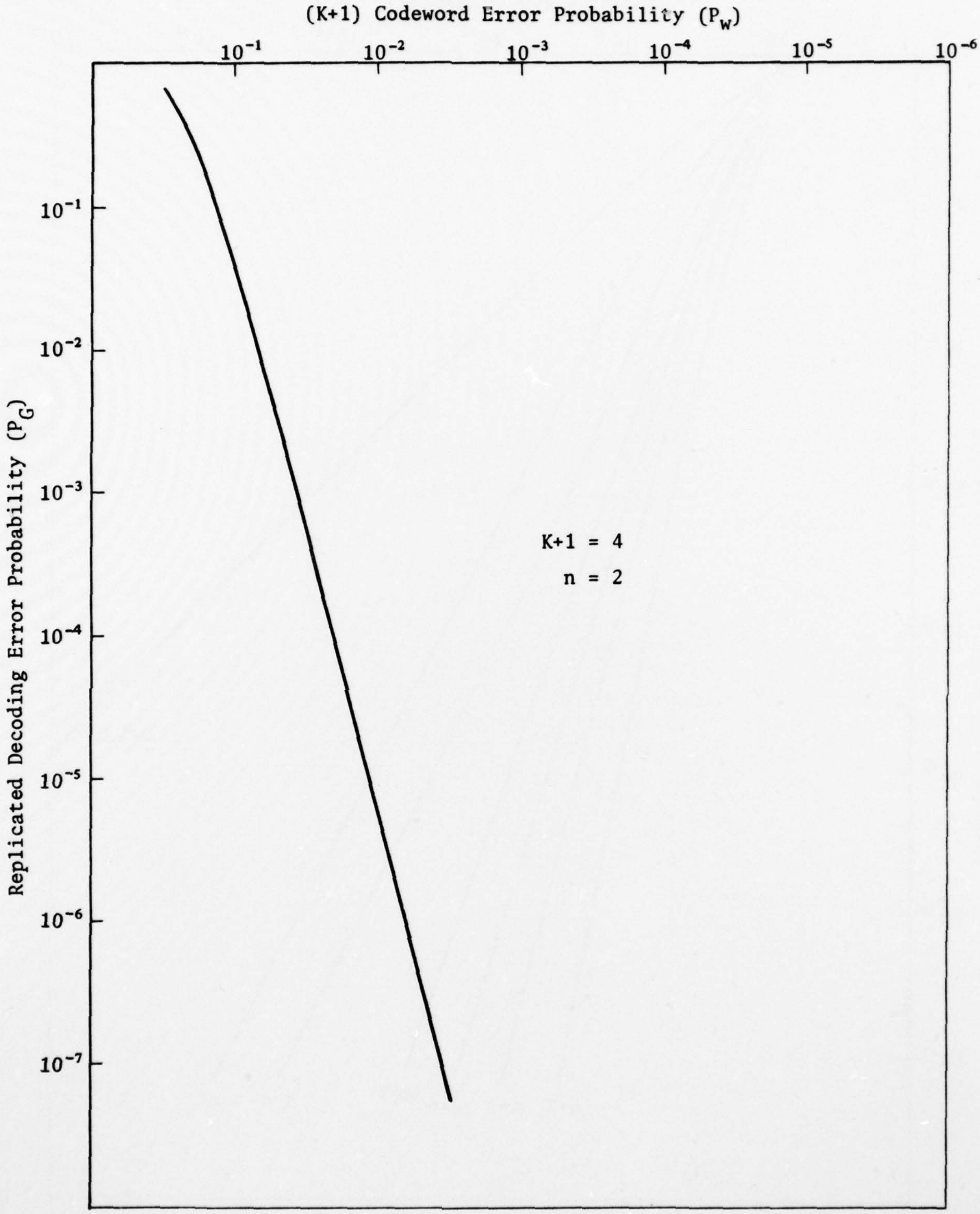


Figure 19. Cascaded Codes. Within Group Replication.
Random Bit Allocation Between Golay Words.

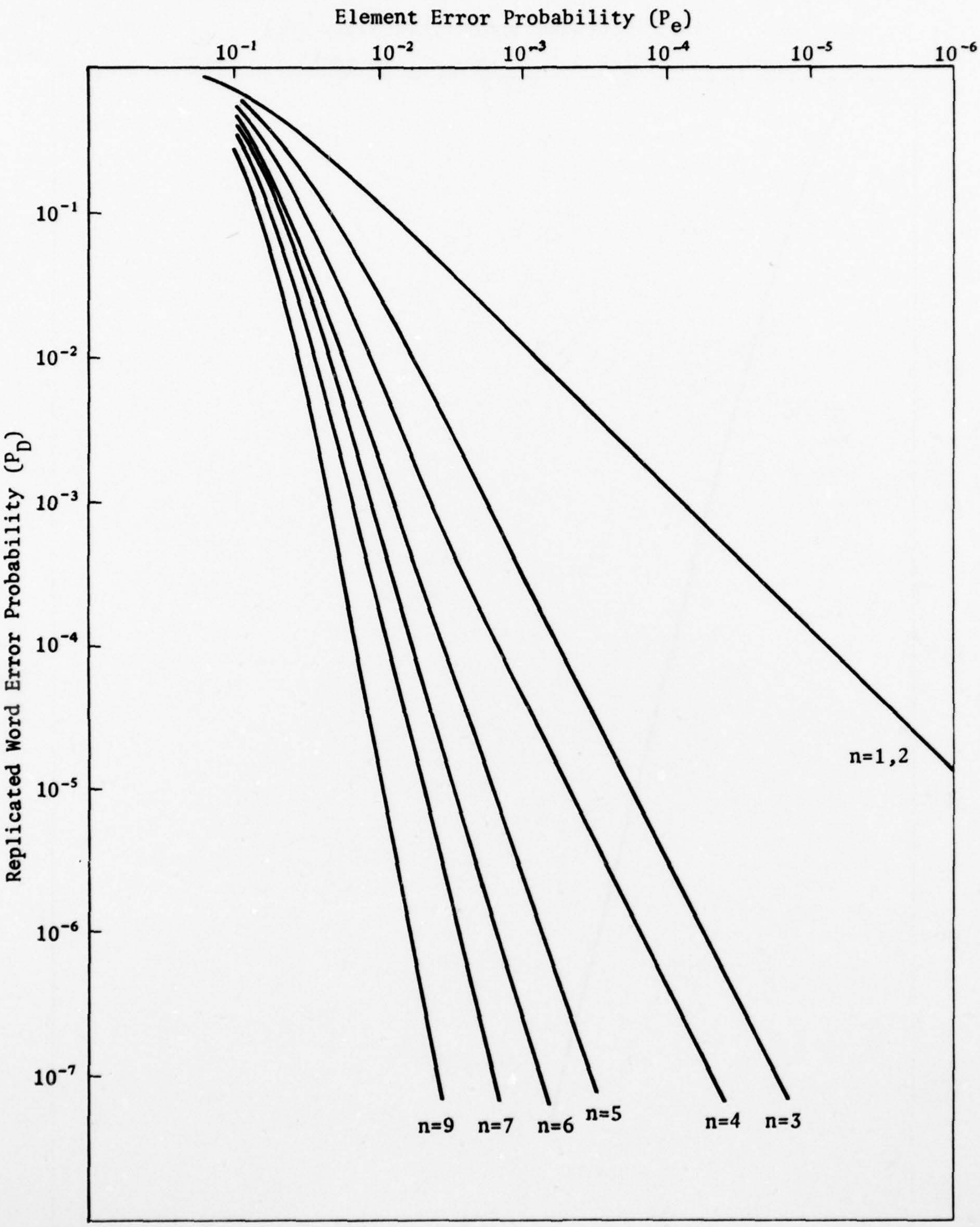


Figure 20. Replication of 12-Bit Data Word

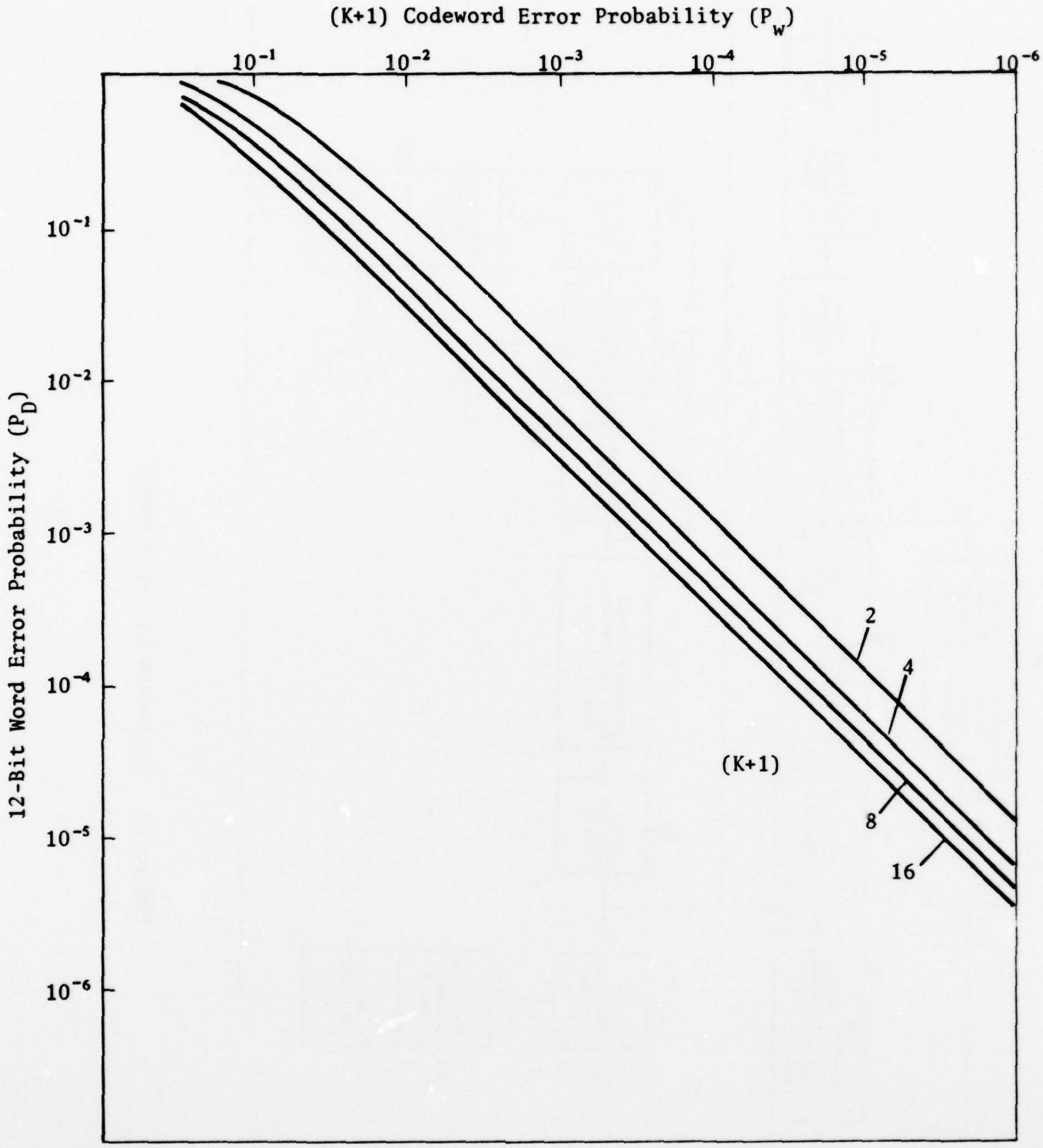


Figure 21. Unreplicated 12-Bit Word Via (K+1) Code

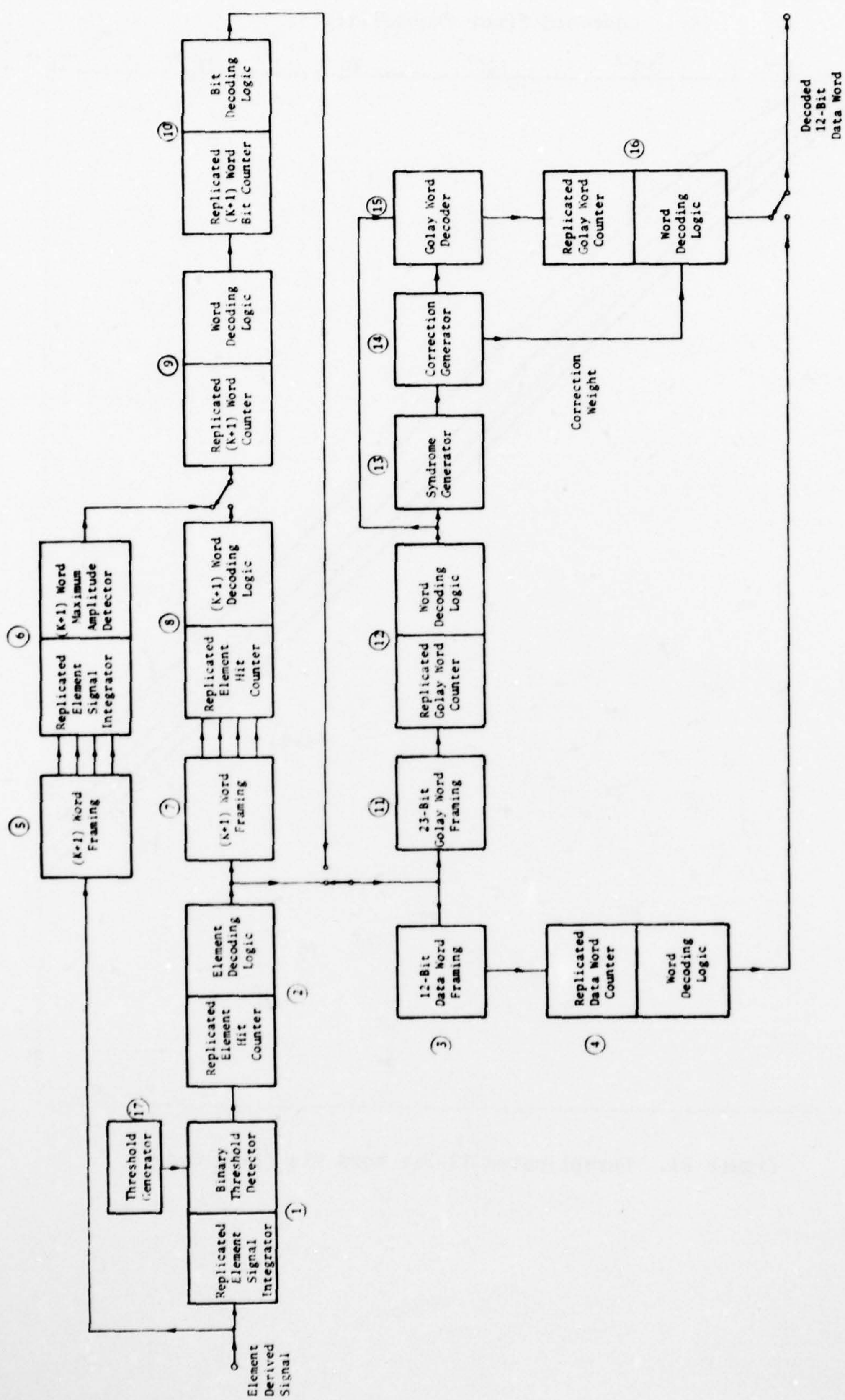


Figure 22. Candidate Coding Schemes

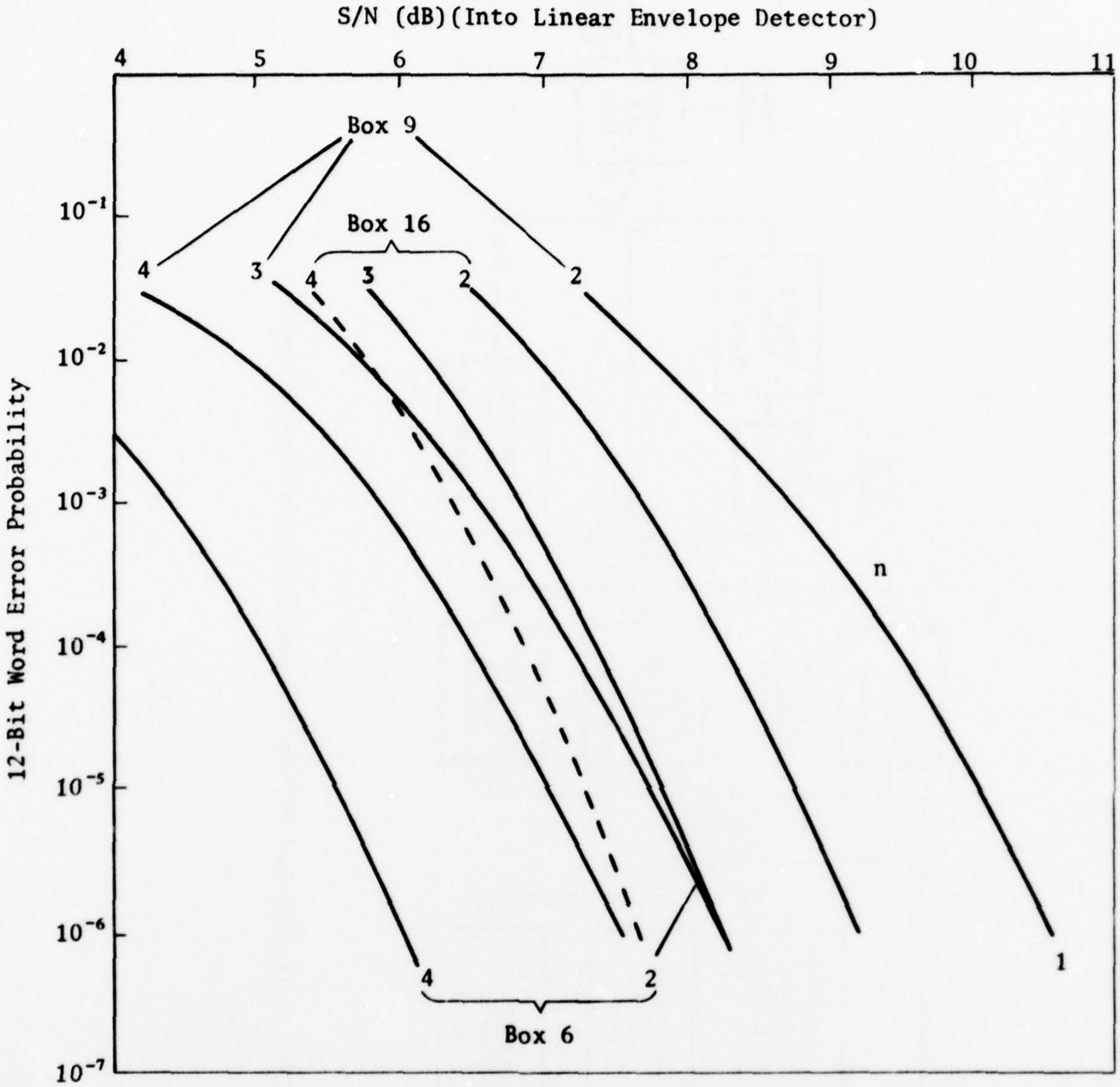


Figure 23. Cascaded Golay and (K+1) Code. (K+1) = 4.
Scheme 6

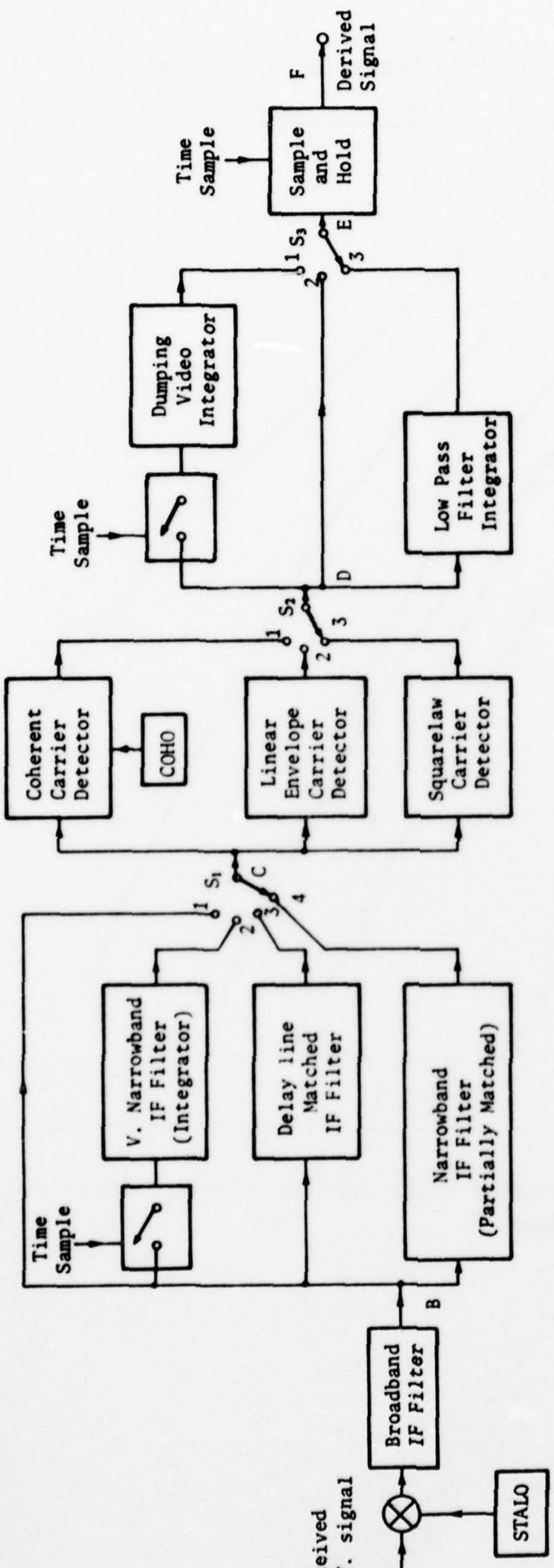


Figure 24. Quasi-Matched Filter Schemes for Individual Elements

DISTRIBUTION

	Copy No.
EXTERNAL	
In United Kingdom	
Defence Scientific and Technical Representative, London	1
In United States of America	
Counsellor, Defence Science, Washington	2
In Australia	
Chief Defence Scientist	3
Director, Joint Intelligence Organisation (DDSTI)	4
Controller, Military Studies and Operational Analysis Division	5
Superintendent, Military Advisers Branch	6
Army Scientific Adviser	7
Navy Scientific Adviser	8
Airforce Scientific Adviser	9
Head, Services Analystical Studies Group	10
Superintendent, Central Studies Establishment	11
Head, Scientific Forecasting Branch	12
Executive Controller, Australian Defence Scientific Service	13
Superintendent, Defence Science Administration Division	14
Head, Laboratory Programmes Branch	15
Defence Information Services Branch (for microfilming)	16
Defence Library, Campbell Park	17
Library, Aeronautical Research Laboratories	18
Library, Materials Research Laboratories	19
Defence Information Services Branch for:	
United Kingdom, Ministry of Defence, Defence Research Information Centre (DRIC)	20
United States, Department of Defense, Defense Documentation Center	21 - 32
Canada, Department of National Defence, Defence Science Information Service	33
New Zealand, Department of Defence	34
Australia National Library	35
INTERNAL	
Director	36
Chief Superintendent, Weapons Research and Development Wing	37
Superintendent, Weapon Systems Division	38
Superintendent, Systems Analysis Division	39
Superintendent, Electronics Division	40

	Copy No.
Superintendent, Communications and Electronic Engineering Division	41
Senior Principal Research Scientist, Radar	42
Senior Principal Research Scientist, Marine	43
Senior Principal Research Scientist, Electronic Warfare	44
Principal Officer, System Development Group	45
Principal Officer, Systems Integration Group	46
Principal Officer, Systems Modelling Group	47
Principal Officer, Target Development Group	48
Principal Officer, Terminal Guidance Systems Group	49
Principal Officer, Tracking and Command Systems Group	50
Principal Officer, Radar and Electronic Tracking Group	51
Principal Officer, Cybernetic Systems Group	52
Principal Officer, Jindalee Group	53
Principal Officer, Radio Systems Group	54
Principal Officer, Electronic Warfare Studies Group	55
Principal Officer, Electronic Warfare Development Group	56
Principal Officer, Ionospheric Studies Group	57
Principal Officer, Target Systems Group	58
Principal Officer, Computing Services Group	59
Principal Officer, Communications Technology Group	60
Dr. K.W. Sarkies, System Development Group	61
Dr. G.B. Brownlie, System Development Group	62
Author	63 - 64
Library, W.R.E.	65 - 70
Spares	71 - 80